

การพัฒนาแบบจำลองการเรียนรู้เชิงลึกนำหน้าเพื่อการรู้จำกิจกรรมของ  
มนุษย์ โดยอาศัยข้อมูลเซนเซอร์จากสมาร์ทโฟน



พรนิภา จันทวงศ์

วิทยานิพนธ์เสนอมหาวิทยาลัยพะเยา เพื่อเป็นส่วนหนึ่งของการศึกษา

หลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์

มิถุนายน 2566

ลิขสิทธิ์เป็นของมหาวิทยาลัยพะเยา

การพัฒนาแบบจำลองการเรียนรู้เชิงลึกนำหน้าเพื่อการรู้จำกิจกรรมของมนุษย์ โดยอาศัย  
ข้อมูลเซนเซอร์จากสมาร์ทโฟน



พรนิภา จันทวงศ์

วิทยานิพนธ์เสนอมหาวิทยาลัยพะเยา เพื่อเป็นส่วนหนึ่งของการศึกษา

หลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์

มิถุนายน 2566

ลิขสิทธิ์เป็นของมหาวิทยาลัยพะเยา

A LIGHTWEIGHT DEEP LEARNING MODEL EFFICIENT HUMAN ACTIVITY RECOGNITION  
USING SMARTPHONE SENSORS



PONNIPA JANTAWONG

A Thesis Submitted to University of Phayao  
in Partial Fulfillment of the Requirements  
for the Master of Engineering Degree in Computer Engineering  
June 2023

Copyright 2023 by University of Phayao

วิทยานิพนธ์

เรื่อง

การพัฒนาแบบจำลองการเรียนรู้เชิงลึกนำหน้าเพื่อการรู้จำกิจกรรมของมนุษย์ โดยอาศัย  
ข้อมูลเซนเซอร์จากสมาร์ตโฟน

ของ พรนิภา จันทวงศ์

ได้รับพิจารณาอนุมัติให้เป็นส่วนหนึ่งของการศึกษา

หลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์

ของมหาวิทยาลัยพะเยา

ประธานกรรมการสอบวิทยานิพนธ์

(ผู้ช่วยศาสตราจารย์ ดร. เกียรติศักดิ์ เตมีย)

ประธานที่ปรึกษาวิทยานิพนธ์

(ผู้ช่วยศาสตราจารย์ ดร. สาทศร เมฆรักขานิช)

อาจารย์บัณฑิตศึกษามหาวิทยาลัยพะเยา

(ผู้ช่วยศาสตราจารย์ ดร. บวรศักดิ์ ศรีสังสิทธิ์สันติ)

อาจารย์บัณฑิตศึกษามหาวิทยาลัยพะเยา

(ดร. นราศักดิ์ บุญเทพ)

คณบดีคณะเทคโนโลยีสารสนเทศและการสื่อสาร

(ผู้ช่วยศาสตราจารย์ ดร. พรเทพ โรจนวสุ)

<b>เรื่อง:</b>	การพัฒนาแบบจำลองการเรียนรู้เชิงลึกน้ำหนักเบาเพื่อการรู้จำกิจกรรมของมนุษย์ โดยอาศัยข้อมูลเซนเซอร์จากสมาร์ตโฟน
<b>ผู้วิจัย:</b>	พรนิภา จันทวงศ์, วิทยานิพนธ์: วศ.ม.(วิศวกรรมคอมพิวเตอร์), มหาวิทยาลัยพะเยา, 2566
<b>อาจารย์ที่ปรึกษา:</b>	ผู้ช่วยศาสตราจารย์ ดร. สาคร เมษรัรักษาวณิช
<b>คำสำคัญ:</b>	การรู้จำกิจกรรมของมนุษย์, การเรียนรู้เชิงลึก, สัญญาณเซ็นเซอร์ที่สวมใส่ได้, โครงข่ายประสาทเทียม, เซ็นเซอร์สมาร์ตโฟน

### บทคัดย่อ

การศึกษาวิจัยครั้งนี้มีวัตถุประสงค์ 1) เพื่อพัฒนาอัลกอริทึมการเรียนรู้เชิงลึกน้ำหนักเบาในการเรียนรู้เชิงลึกสำหรับการเรียนรู้กิจกรรมในชีวิตประจำวันของมนุษย์ที่มีประสิทธิภาพ 2) เพื่อนำผลลัพธ์ที่ได้ไปพัฒนานวัตกรรมในการดูแลมนุษย์ในด้านต่าง ๆ โดยได้กำหนดขอบเขตของการวิจัย 3 ขั้นตอนดังนี้ การประมวลผลเบื้องต้น ขั้นตอนการเตรียมข้อมูล และขั้นตอนการเรียนรู้กิจกรรมในชีวิตประจำวันของมนุษย์ ในการทดลองวิจัยครั้งนี้ได้ใช้ชุดข้อมูล SPARS9x ซึ่งเป็นชุดข้อมูลสาธารณะ เครื่องมือที่ใช้ในการวิจัยได้แก่ โคแลปโปร และใช้ภาษาไพธอนในการพัฒนา การทดลองกับแบบจำลอง 6 แบบจำลอง ได้แก่ VGG16, ResNet18, PyramidNet18, Inception-V3, Xception และ EffcientNet-B0 ผลการทดลองพบว่า PyramidNet18 มีความแม่นยำสูงสุดที่ 99.15% และ F1-score 99.15% และพบว่าแบบจำลองแบบไฮบริดที่พัฒนาร่วมกับโครงข่ายประสาทเทียมสามารถจำแนกและดึงคุณสมบัติพื้นฐานออกมาได้เป็นอย่างดี และยังแสดงให้เห็นถึงความสามารถในการคำนวณอีกด้วย

**Title:** A LIGHTWEIGHT DEEP LEARNING MODEL EFFICIENT HUMAN ACTIVITY RECOGNITION USING SMARTPHONE SENSORS

**Author:** PONNIPA JANTAWONG, Thesis: M.En. (Computer Engineering), University of Phayao, 2023

**Advisor:** Assistant Professor Dr. Sakorn Mekruksavanich

**Keywords:** human activity recognition, deep learning, wearable sensor signal, convolutional neural network, smartphone sensors

### ABSTRACT

The objective of this research was two-fold: 1) to develop lightweight deep learning algorithms that facilitate efficient learning of human daily activities, and 2) to utilize the resulting algorithm model to contribute to the development of innovative solutions that support human care in various fields. The research focused on three main areas: data pre-processing methods, data generation methods, and model training methods for classifying human daily activities. To conduct the experiments, a publicly available dataset called SPARS9x was utilized. The research employed Colab Pro as the tool and utilized the Python language for development purposes. The dataset included six different models, namely VGG16, ResNet18, PyramidNet18, Inception-V3, Xception, and EfficientNet-B0. The experimental findings revealed that PyramidNet18 achieved the highest accuracy of 99.15% and an F1-score of 99.15% when tested on the SPARS9x dataset. Additionally, it was discovered that hybrid models deployed with convolutional neural networks not only provided outstanding results but also demonstrated computational efficiency.



## กิตติกรรมประกาศ

การศึกษาวิจัยเรื่องการพัฒนาแบบจำลองการเรียนรู้เชิงลึกน้ำหนักเบาเพื่อการรู้จำกิจกรรมของมนุษย์โดยอาศัยข้อมูลเซนเซอร์จากสมาร์ตโฟนฉบับนี้สำเร็จลุล่วงลงได้ด้วยความรู้จาก ผู้ช่วยศาสตราจารย์ ดร.สาคร เมฆรักษาวิช และผู้ช่วยศาสตราจารย์ ดร.ศกยภพ ประเวทจิตร อาจารย์ที่ปรึกษาที่ให้ความกรุณาแนะนำ ตรวจสอบแก้ไขปัญหา ข้อเสนอแนะ ข้อบกพร่องต่าง ๆ และติดตามความก้าวหน้าของงานวิจัยฉบับนี้จนเสร็จสมบูรณ์

ผู้วิจัยขอขอบพระคุณอาจารย์ประจำสาขาวิชาวิศวกรรมคอมพิวเตอร์ เทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา ทุกท่านที่ได้ประสิทธิ์ ประสาทวิชาความรู้ทั้งนอกห้องเรียน และในห้องเรียน เพื่อให้ได้รับความรู้มากที่สุดในการเรียน จนทำให้การศึกษาวิจัยครั้งนี้ สมบูรณ์

พรนิภา จันทวงศ์



## สารบัญ

	หน้า
บทคัดย่อภาษาไทย .....	ง
บทคัดย่อภาษาอังกฤษ .....	จ
กิตติกรรมประกาศ .....	ฉ
สารบัญ .....	ช
สารบัญตาราง .....	ฅ
สารบัญภาพ.....	ญ
บทที่ 1 บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา .....	1
วัตถุประสงค์ของการวิจัย.....	2
ขอบเขตของการวิจัย .....	2
ประโยชน์ที่คาดว่าจะได้รับ .....	2
ขั้นตอนของการดำเนินงานวิจัย.....	3
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง .....	4
แนวคิดและทฤษฎีเบื้องต้นที่เกี่ยวข้องของงานวิจัย.....	4
ไบโอเมตริกซ์ .....	4
การรู้จำกิจกรรมของมนุษย์ (Human Activity Recognition : HAR).....	4
เซนเซอร์ที่เกี่ยวข้องในการรู้จำกิจกรรมของมนุษย์.....	4
ชุดข้อมูลสาธารณะสำหรับกิจกรรมของมนุษย์.....	9
การเรียนรู้เชิงลึก .....	11
มาตรวัดประสิทธิภาพของแบบจำลองแบบหมวดหมู่.....	17
งานวิจัยที่เกี่ยวข้อง .....	19

บทที่ 3 วิธีดำเนินการวิจัย.....	39
การเก็บข้อมูล (collection).....	40
การประมวลผลเบื้องต้น (processing).....	45
การเตรียมข้อมูล (data generation) .....	49
การฝึกฝนแบบจำลอง และการจัดหมวดหมู่กิจกรรมของมนุษย์ (training model) .....	50
บทที่ 4 ผลการวิจัย.....	60
บทที่ 5 สรุปผลการวิจัย.....	69
บรรณานุกรม.....	71
ประวัติผู้วิจัย .....	76



## สารบัญตาราง

	หน้า
ตาราง 1 การแบ่งชุดข้อมูลการฝึกอบรมและทดสอบของชุดข้อมูล UCI .....	21
ตาราง 2 ประสิทธิภาพการทดสอบของ 6 กิจกรรม.....	23
ตาราง 3 อธิบายชุดข้อมูล WISDM .....	27
ตาราง 4 รายละเอียดของชุดข้อมูลที่ใช้ในงานวิจัย Xia et al. (2020).....	31
ตาราง 5 ผลลัพธ์จากจำลอง iSPLInception กับแบบจำลองอื่น ๆ จากชุดข้อมูล 4 ชุด .....	37
ตาราง 6 กิจกรรมในชุดข้อมูล PAMAP2 .....	41
ตาราง 7 กิจกรรมการเคลื่อนไหวในชีวิตประจำวันของชุดข้อมูลสาธารณะ w-HAR .....	42
ตาราง 8 เวลาในการทำกิจกรรมการเคลื่อนไหวในชีวิตประจำวันของชุดข้อมูลสาธารณะ w-HAR	43
ตาราง 9 โครงสร้างแบบจำลอง 6 แบบกับชุดข้อมูล PAMAP2 .....	61
ตาราง 10 ผลการเปรียบเทียบกับแบบจำลองทั้ง 7 แบบกับชุดข้อมูล PAMAP2.....	61
ตาราง 11 ผลการเปรียบเทียบแต่ละกิจกรรมกับแบบจำลองทั้ง 6 แบบกับชุดข้อมูล PAMAP2.....	63
ตาราง 12 ผลการเปรียบเทียบแบบจำลองโครงข่ายประสาทเทียม 6 แบบ กับชุดข้อมูล SPARS9x .65	65
ตาราง 13 ผลการเปรียบเทียบกับแบบจำลองทั้ง 5 แบบกับชุดข้อมูล w-HAR.....	66
ตาราง 14 จำนวนพารามิเตอร์ของแบบจำลองทั้ง 5 แบบกับชุดข้อมูล w-HAR .....	67
ตาราง 15 ผลการเปรียบเทียบกับแบบจำลอง CNN LSTM CNN-LSTM กับชุดข้อมูล MEX ใช้ เซ็นเซอร์วัดความเร่งบนตำแหน่งคันท้าย.....	67
ตาราง 16 ผลการเปรียบเทียบกับแบบจำลอง CNN LSTM CNN-LSTM กับชุดข้อมูล MEX ใช้ เซ็นเซอร์วัดความเร่งบนตำแหน่งข้อมือ .....	68

## สารบัญภาพ

	หน้า
ภาพ 1 หลักการพื้นฐานของกฎข้อที่สองของนิวตันในรูปแบบสองมิติ .....	5
ภาพ 2 การออกแบบของเซนเซอร์วัดความเร่ง (Countryman, 2014) .....	6
ภาพ 3 หลักการพื้นฐานของเซนเซอร์วัดการหมุน (Grahn, 2017) .....	6
ภาพ 4 ลักษณะการทำงานของเซนเซอร์วัดการหมุน (ทีณะ, 2560) .....	7
ภาพ 5 โครงสร้างของเซนเซอร์ GMR (Mohamadabadi, 2013) .....	8
ภาพ 6 การรู้จักกิจกรรมของมนุษย์จากเซนเซอร์โดยใช้การเรียนรู้เชิงลึก .....	12
ภาพ 7 การทำงานของ CNN .....	13
ภาพ 8 การทำงานของ RNN .....	13
ภาพ 9 การทำงานของ LSTM .....	14
ภาพ 10 การทำงานของ GRU .....	15
ภาพ 11 การทำงานของ Bi-LSTM .....	16
ภาพ 12 การทำงานของ Bi-GRU .....	16
ภาพ 13 การทำงานของ Random Forest .....	17
ภาพ 14 สถาปัตยกรรมของแบบจำลอง CNN (Ignatov Andrey, 2018) .....	19
ภาพ 15 การทำงานของโครงข่ายประสาทเทียม Coarse-fine Carlos .....	20
ภาพ 16 ค่าความสูญเสียของชุดฝึกฝน Coarse-Fine .....	22
ภาพ 17 ค่าความแม่นยำของชุดทดสอบ Coarse-Fine .....	22
ภาพ 18 ขั้นตอนพื้นฐานการเรียนรู้กิจกรรมของมนุษย์ .....	23
ภาพ 19 โหนดของ RNN .....	24
ภาพ 20 เซลล์หน่วยความจำของ LSTM .....	25

ภาพ 21 การทำงานของ Lightweight RNN–LSTM Preeti et al. (2020) .....	26
ภาพ 22 โครงสร้างในการอัปเดตเซลล์ LSTM.....	28
ภาพ 23 โครงสร้างของ LSTM.....	28
ภาพ 24 โครงสร้างของแบบจำลอง H–LSTM.....	29
ภาพ 25 การทำงานของแบบจำลอง LSTM–CNN Kun et al. (2020).....	32
ภาพ 26 การรวมชั้นเครือข่าย Convolutional และ Long Short–Term Memory .....	34
ภาพ 27 โมดูล Inception ของแบบจำลอง iSPLInception.....	35
ภาพ 28 แบบจำลอง iSPLInception.....	36
ภาพ 29 แสดงขั้นตอนกระบวนการเรียนรู้จำกิจกรรมในชีวิตประจำวันของมนุษย์ .....	39
ภาพ 30 แสดงภาพรวมการเก็บรวบรวมข้อมูล.....	40
ภาพ 31 แสดงการใช้งาน pandas อ่านไฟล์นามสกุล csv.....	44
ภาพ 32 แสดงภาพรวมการประมวลผลข้อมูล.....	45
ภาพ 33 แสดงตัวอย่างการใช้งาน Butterworth แบบ Low–pass filter .....	46
ภาพ 34 แสดงตัวอย่างการใช้งาน Min–Max Normalization.....	47
ภาพ 35 แสดงตัวอย่างการใช้งาน Standardization.....	47
ภาพ 36 แสดงตัวอย่างการแบ่งส่วนข้อมูลที่ทับซ้อนกัน 50% .....	48
ภาพ 37 แสดงภาพรวมการเตรียมข้อมูล .....	49
ภาพ 38 การแบ่งชุดข้อมูลสำหรับการฝึกฝน การทดสอบ .....	49
ภาพ 39 ตัวอย่างการใช้งาน K–Fold Cross Validation .....	50
ภาพ 40 การฝึกฝนแบบจำลอง และจัดหมวดหมู่กิจกรรมของมนุษย์.....	50
ภาพ 41 โครงสร้างหลักของโครงข่ายประสาทเทียมแบบคอนโวลูชันที่ใช้ในงานวิจัยนี้.....	51
ภาพ 42 โครงสร้างแบบจำลอง VGG16.....	52
ภาพ 43 โครงสร้างแบบจำลอง RESNET18.....	53

ภาพ 44 โครงสร้างแบบจำลอง PyramidNet18 .....54

ภาพ 45 โครงสร้างแบบจำลอง Inception-V3 .....55

ภาพ 46 โครงสร้างแบบจำลอง Xception.....56

ภาพ 47 โครงสร้างแบบจำลอง EfficientNet-B0.....57

ภาพ 48 โครงสร้างแบบจำลอง LSTM.....57

ภาพ 49 โครงสร้างแบบจำลอง GRU.....58

ภาพ 50 โครงสร้างแบบจำลอง Bi-LSTM.....58

ภาพ 51 โครงสร้างแบบจำลอง Bi-GRU.....59

ภาพ 52 แสดงภาพรวมการทำงานสำหรับเรียนรู้จำกิจกรรมในชีวิตประจำวันของมนุษย์ .....60

ภาพ 53 โครงสร้างแบบจำลอง InceptionTime .....62

ภาพ 54 fix-width sliding window .....64

ภาพ 55 โครงสร้างแบบจำลอง Lightweight (Z. Zhongkai, 2022) .....65

ภาพ 56 จำนวนพารามิเตอร์ของแบบจำลอง Lightweight (Z. Zhongkai, 2022).....66



## บทที่ 1

### บทนำ

#### ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันโครงสร้างประชากรของประเทศไทยได้เปลี่ยนแปลงไปจากเดิมลักษณะของประชากรกลุ่มผู้สูงอายุมีแนวโน้มเพิ่มมากขึ้น จากการศึกษาขนาดและแนวโน้มของประชากรผู้สูงอายุ จากสำมะโนประชากร ประเทศไทยพบว่าตั้งแต่ปี พ.ศ. 2552 ประเทศไทยก้าวเข้าสู่ยุคสังคมผู้สูงอายุ (Aging Society) อย่างสมบูรณ์ มีจำนวนผู้สูงอายุมากกว่าร้อยละ 10 ของประชากรทั้งหมด (Shryock, 2004) ซึ่งปัญหาที่เกิดขึ้นกับผู้สูงอายุส่วนใหญ่ล้วนเป็นปัญหาทางด้านสุขภาพ ทั้งทางด้านสุขภาพกายและสุขภาพจิต เพราะเป็นวัยที่ต้องเผชิญกับการเปลี่ยนแปลงในทางเสื่อมอย่างหลีกเลี่ยงไม่ได้ ส่งผลให้ผู้สูงอายุมีบทบาททางสังคมลดลงและเปลี่ยนไปจากเดิม ทั้งการดำเนินชีวิต อาชีพ ลักษณะอารมณ์ ความสัมพันธ์กับบุคคลอื่น ทั้งในครอบครัวและในสังคม โดยกลุ่มผู้สูงอายุมีภาวะเสี่ยงต่อการเจ็บป่วย ด้วยโรคเรื้อรัง หรือโรคประจำตัวสูงกว่าทุกกลุ่มอายุ เมื่อพิจารณาถึงต้นเหตุ ที่ทำให้เกิดโรคเหล่านั้น จะเห็นได้ว่า ส่วนใหญ่มีสาเหตุมาจากพฤติกรรมกรดำเนินชีวิตหรือการดูแลสุขภาพตนเอง ซึ่งสามารถหลีกเลี่ยงหรือบรรเทาความรุนแรงของโรคได้ หากผู้สูงอายุได้รับดูแลสุขภาพให้ถูกต้องและเหมาะสม รวมทั้งจะต้องมีการเตรียมความพร้อมในระบบบริการสังคมและสุขภาพที่ดี

ในยุคศตวรรษที่ 21 เทคโนโลยีมีวิวัฒนาการแบบก้าวกระโดด การประยุกต์ใช้เทคโนโลยีในด้านต่างๆ ถูกพัฒนาขึ้นอย่างต่อเนื่อง ส่งผลให้มีกรอบแนวคิดที่จะนำเอาเทคโนโลยีสารสนเทศเข้ามาพัฒนาการรู้จักกิจกรรมของมนุษย์ เพื่อตอบโจทย์สำหรับมนุษย์ โดยเฉพาะผู้สูงอายุ โดยปัจจุบันได้มีการพัฒนาอัลกอริทึมการเรียนรู้จำการทำกิจกรรมประจำวันของมนุษย์ในด้านต่าง ๆ อาทิเช่น การดูแลสุขภาพ การออกกำลังกาย รวมถึงการวิเคราะห์การเคลื่อนไหวของมนุษย์อีกด้วย โดยกิจกรรมต่าง ๆ ในชีวิตประจำวันของมนุษย์ที่ได้มีการนำมาพัฒนาแบบจำลองจะมีลักษณะของกิจกรรมที่ไม่ซับซ้อนของมนุษย์ (Simple Activity) เช่น การเดิน การวิ่ง ฯลฯ และกิจกรรมที่มีความซับซ้อนของมนุษย์ (Complex Activity) เช่น การอาบน้ำ การเก็บของ การทำอาหาร ฯลฯ โดยใช้อัลกอริทึมของการเรียนรู้เชิงลึก (Deep learning : DL) การเรียนรู้เชิงลึกเป็นการเรียนรู้ของเครื่องคอมพิวเตอร์ มีโครงสร้างที่มีลักษณะ

เหมือนเซลล์ประสาทหลาย ๆ เซลล์เชื่อมต่อกัน เรียกว่าโหนด เส้นประสาท และเซลล์ต่าง ๆ จะเชื่อมต่อกันเป็นชั้น ๆ โดยแบ่งเป็น 3 ชั้น ได้แก่ ชั้นนำเข้าข้อมูล (Input Layer) ชั้นที่เป็นระบบเซลล์ประสาทในสมอง (Hidden Layer) และชั้นข้อมูลขาออก (Output Layer) การเรียนรู้เชิงลึกจะฝึกฝน เรียนรู้ด้วยตนเอง เพื่อให้สามารถทำงานได้เหมือนมนุษย์ เช่น การระบุรูปภาพ การคาดการณ์ หรือการจดจำเสียง ซึ่งถูกพัฒนา และปรับปรุงอย่างต่อเนื่อง ทั้งการปรับปรุงเพื่อเพิ่มประสิทธิภาพของอัลกอริทึมการปรับปรุงในเรื่องความถูกต้องของอัลกอริทึมให้มีความแม่นยำมากขึ้น

จากที่กล่าวมาข้างต้น ทางผู้วิจัยจึงมีกรอบแนวคิดในการพัฒนาอัลกอริทึมการเรียนรู้เชิงลึกที่น้ำหนักเบาสำหรับการเรียนรู้จากกิจกรรมของมนุษย์ให้มีความแม่นยำ และมีการทำงานที่รวดเร็ว ต่อไป

### วัตถุประสงค์ของการวิจัย

1. เพื่อทำการพัฒนาอัลกอริทึมที่น้ำหนักเบาในการเรียนรู้เชิงลึกสำหรับการเรียนรู้จากกิจกรรมในชีวิตประจำวันของมนุษย์ที่มีประสิทธิภาพ

### ขอบเขตของการวิจัย

นำเสนอแบบจำลองสำหรับขั้นตอนการประมวลผลเบื้องต้น (data preprocessing) ขั้นตอนการเตรียมข้อมูล (data generation) และขั้นตอนการเรียนรู้จากกิจกรรมในชีวิตประจำวันของมนุษย์ (model training) ที่มีประสิทธิภาพ

### ประโยชน์ที่คาดว่าจะได้รับ

1. มีความรู้ความเข้าใจอัลกอริทึมในขั้นตอนการประมวลผลเบื้องต้น (data preprocessing) ขั้นตอนการเตรียมข้อมูล (data generation) และขั้นตอนการเรียนรู้จากกิจกรรมในชีวิตประจำวันของมนุษย์ (model training)
2. ได้แบบจำลองใหม่สำหรับขั้นตอนการประมวลผลเบื้องต้น (data preprocessing) ขั้นตอนการเตรียมข้อมูล (data generation) และขั้นตอนการเรียนรู้จากกิจกรรมในชีวิตประจำวันของมนุษย์ (model training)

3. ใ้งานวิจัยที่เป็นต้นแบบสำหรับการเรียนรู้เชิงลึกในการเรียนรู้จำกิจกรรมในชีวิตประจำวันของมนุษย์

#### ขั้นตอนของการดำเนินงานวิจัย

1. ศึกษางานวิจัยทางด้านการเรียนรู้จำกิจกรรมในชีวิตประจำวันของมนุษย์
2. ศึกษาทฤษฎีที่เกี่ยวข้องกับการพัฒนาแบบจำลองการเรียนรู้จำกิจกรรมในชีวิตประจำวันของมนุษย์ที่มีประสิทธิภาพ
3. ออกแบบแบบจำลองใหม่
4. พัฒนาอัลกอริทึมตามรูปแบบของแบบจำลองที่ได้ออกแบบไว้ โดยใช้โคแลบโปร (colab pro)
5. จัดเตรียมข้อมูลสำหรับการทดลองวิจัยแล้วดำเนินการทดลองเพื่อวัดประสิทธิภาพของแบบจำลองที่นำเสนอ
6. สรุปผลการทดลอง
7. จัดทำรายงานการวิจัย

## บทที่ 2

### เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้ ผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง และได้นำเสนอตามหัวข้อดังต่อไปนี้ แนวคิดและทฤษฎีเบื้องต้นที่เกี่ยวข้องของในงานวิจัย และงานวิจัยที่เกี่ยวข้อง **แนวคิดและทฤษฎีเบื้องต้นที่เกี่ยวข้องของงานวิจัย**

#### ไบโอเมตริกซ์

ไบโอเมตริกซ์ คือเทคโนโลยีที่ตรวจวัดคุณลักษณะหรือลักษณะทางพฤติกรรมของมนุษย์ โดยนำมาใช้ในการระบุตัวบุคคลนั้น ๆ หรือกล่าวว่าเป็นเทคโนโลยีสำหรับยืนยันตัวบุคคล ไบโอเมตริกซ์แบ่งประเภทเป็น 2 ประเภท คือ ลักษณะทางกายภาพ (Physiological Biometrics) คือข้อมูลที่ปรากฏอยู่ภายในร่างกาย หรืออวัยวะต่าง ๆ ของบุคคล เช่น ลายนิ้วมือ ลายม่านตา ลายนิ้วมือ ลักษณะใบหน้า ลักษณะของมือ ลักษณะของนิ้วมือ ลักษณะใบหู เสียงพูด เป็นต้น และลักษณะทางพฤติกรรม (Behavioural Biometrics) คือข้อมูลที่มาจากพฤติกรรม หรือการกระทำที่เป็นเอกลักษณ์ของบุคคล เช่น การกดแป้นพิมพ์ การเดิน การลงลายมือชื่อ เสียงพูด เป็นต้น (เอกภพ สกุลกิจกาญจน์, 2552)

การรู้จำกิจกรรมของมนุษย์ (Human Activity Recognition : HAR)

การรู้จำกิจกรรมของมนุษย์ เป็นการวิเคราะห์ข้อมูลการเคลื่อนไหวจากเซนเซอร์ เช่น เซนเซอร์วัดความเร่ง เซนเซอร์วัดการหมุน เซนเซอร์วัดค่าความเข้มสนามแม่เหล็ก บลูทูธ และอื่น ๆ ซึ่งเป็นการนำข้อมูลจากเซนเซอร์มาวิเคราะห์ และประมวลผล ถูกนำมาใช้กันอย่างแพร่หลาย เนื่องจากได้รับการปกป้องความเป็นส่วนตัว และมีความปลอดภัยของข้อมูลสูง (Wang และคณะ, 2019)

#### เซนเซอร์ที่เกี่ยวข้องในการรู้จำกิจกรรมของมนุษย์

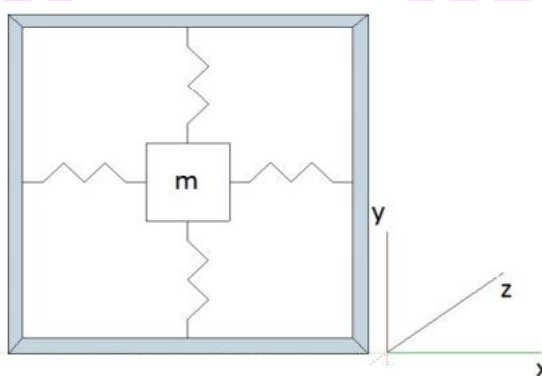
เซนเซอร์มีหลายประเภท แต่ในที่นี้จะกล่าวถึงเซนเซอร์ของสมาร์ทโฟนที่ใช้ในการรู้จำกิจกรรมของมนุษย์ เซนเซอร์ของสมาร์ทโฟนแต่ละแบบมีลักษณะของข้อมูลที่สามารถนำมาใช้ในการวิเคราะห์แตกต่างกันตามรุ่น และผู้จัดทำหน้า

1. เซนเซอร์วัดความเร่ง (Accelerometer Sensor) เซนเซอร์วัดความเร่ง เป็นมาตรฐานวัดความเร่งแบบอัตโนมัติ ใช้ตรวจจับ และวัดค่าการสั่นสะเทือน สามารถใช้วัดแรงสั่นสะเทือนใน

รถยนต์ เครื่องยนต์ สิ่งปลูกสร้าง และการติดตั้งระบบรักษาความปลอดภัย นอกจากนี้ยังสามารถใช้ในการวัดอุปกรณ์ไฟฟ้า เช่น เกมสามมิติ เม้าส์คอมพิวเตอร์ โทรศัพท์ การวัดแรงสั่นสะเทือนจากแผ่นดินไหว และสามารถนำไปประยุกต์ใช้ในงานทางด้านมัลติมีเดีย เช่น การทำวิดีโอสำหรับการเลือกดูเนื้อหาตามที่ต้องการ (Video on demand) ซึ่งเป็นวิดีโอการเคลื่อนไหวแบบสามมิติ และวัตถุสามมิติที่สามารถเปลี่ยนแปลงภาพได้ (Selany, Nasution และ Purboyo, 2015; Sugara Nashar Luthfi, Purboyo Tito Waluyo และ Prasasti Anggunmeka Luhur, 2018)

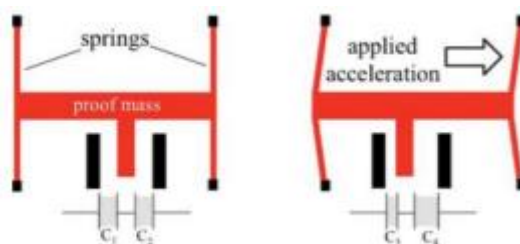
หลักการพื้นฐานของเซนเซอร์วัดความเร่งบนสมาร์ตโฟน เป็นระบบไฟฟ้าเชิงกลจุลภาค หรือเมเนส์ (micro electro mechanical systems : MEMS) แสดงตามกระบวนการทำงานกฎข้อที่สองของนิวตัน  $9.8 \text{ m/s}^2$  ดังภาพ 1

$$F = k \times x = m \times a \quad (1)$$



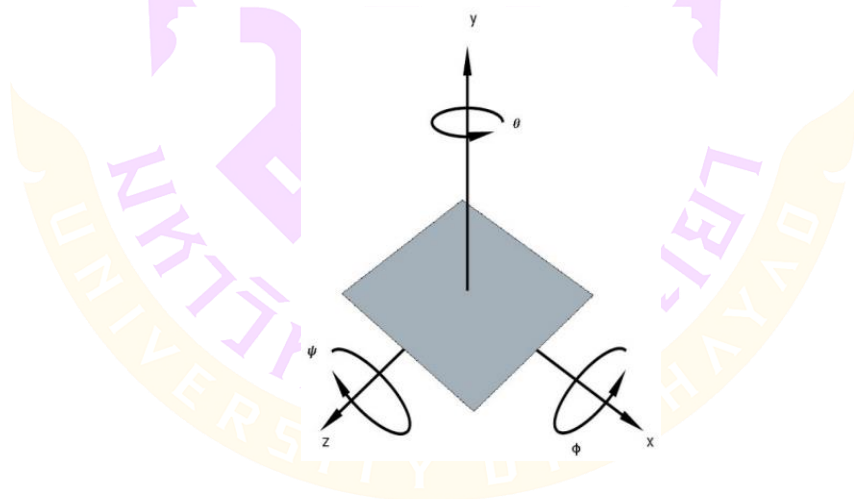
ภาพ 1 หลักการพื้นฐานของกฎข้อที่สองของนิวตันในรูปแบบสองมิติ (Grahn, 2017)

เซนเซอร์วัดความเร่งบนสมาร์ตโฟน เป็นมวลที่เคลื่อนที่ไปมา (seismic) โดยสามารถเคลื่อนที่แบบอิสระไปในทิศทางใดทิศทางหนึ่ง และจะติดตั้งบนเกลียวสปริง ดังภาพ 2



ภาพ 2 การออกแบบของเซนเซอร์วัดความเร่ง (Countryman, 2014)

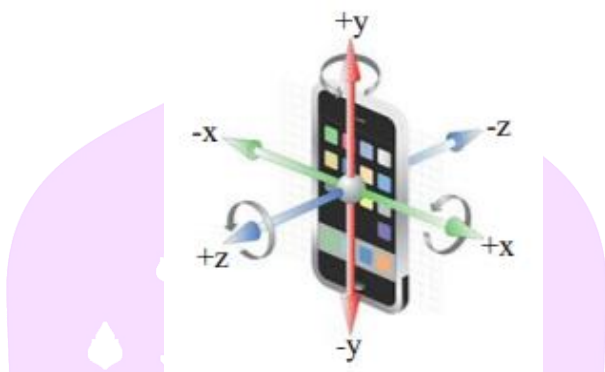
2. เซนเซอร์วัดการหมุน (Gyroscope Sensor) เซนเซอร์วัดการหมุน เป็นเซนเซอร์วัดความเร็วเชิงมุมใช้สำหรับตรวจจับลักษณะการหมุนของสมาร์ทโฟน ใช้การตรวจจับแบบ 3 แกน (3-Axis) เช่นเดียวกับกับเซนเซอร์วัดความเร่ง โดยเซนเซอร์วัดการหมุนจะมีความถูกต้องและแม่นยำมากกว่าเซนเซอร์วัดความเร่ง เช่น การควบคุมการเล่นเกมส์ต่าง ๆ ที่ต้องอาศัยการเคลื่อนไหวในหลาย ๆ ทิศทาง เช่น เกมส์แข่งรถทั้งหลาย การควบคุมอาจจะไม่แม่นยำหากใช้เซนเซอร์วัดความเร่งเพียงอย่างเดียว จึงต้องมีเซนเซอร์วัดการหมุนมาเสริมเพื่อให้การควบคุมมีความถูกต้อง และแม่นยำมากขึ้น (เฉลิมเกียรติ สอนเสียม 2560)



ภาพ 3 หลักการพื้นฐานของเซนเซอร์วัดการหมุน (Grahn, 2017)

เซนเซอร์วัดการหมุนได้รับความนิยมเป็นอย่างมากในการใช้งานต่าง ๆ เช่น การเล่นเกม ระบบป้องกันภาพสั่นไหว และการนำทาง โดยเซนเซอร์นี้ยังประหยัดพลังงาน แต่ยังไม่ครอบคลุมกับการใช้งานที่ต้องการความเสถียรในระยะยาว (Eminoglu, 2017)

ในเชิงพาณิชย์ได้มีการนำเซนเซอร์วัดการหมุนแบบ 3 แกนมาใช้งาน เพื่อลดต้นทุนและขนาด เทคนิคในปัจจุบันไม่อนุญาตให้มีโครงสร้างสมมาตรสำหรับอุปกรณ์นอกเครื่องบิน ดังนั้นมักใช้โครงสร้างที่ไม่สมมาตรเพื่อใช้วัดแกนทั้ง 3 แกน (Prandi และคณะ, 2011)



ภาพ 4 ลักษณะการทำงานของเซนเซอร์วัดการหมุน (ทีชะ, 2560)

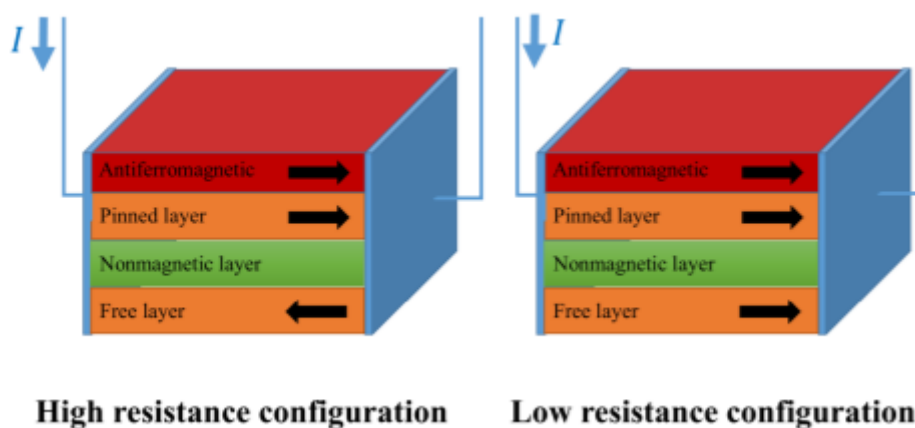
จากภาพ 4 แสดงลักษณะการทำงานของเซนเซอร์วัดการหมุน โดยเซนเซอร์จะมีการเปลี่ยนค่าตามค่าของมุมต่าง ๆ โดยใช้ปรากฏการณ์ของโคริโอลิส (Coriolis Effect) เป็นการใช้อุปกรณ์เคลื่อนที่ไปในทิศทางใดทิศทางหนึ่งด้วยความเร็วเฉพาะของมวล จากนั้นเมื่อเกิดการเปลี่ยนมุมจะทำให้การกระจัดเกิดการเปลี่ยนแปลงค่าความจุที่วัดได้ ผลที่ได้จะสัมพันธ์กับอัตราการหมุนเชิงมุมโดยเฉพาะ (ทีชะ, 2560)

3. เซนเซอร์วัดค่าความเข้มสนามแม่เหล็ก (Magnetometer) เซนเซอร์วัดค่าความเข้มสนามแม่เหล็ก มีด้วยกันหลายประเภท ประเภทที่ใช้กันทั่วไปมากที่สุดในสมาร์ทโฟนมีพื้นฐานมาจากเซนเซอร์ AMR และ GMR ดังนี้

3.1 Anisotropic magnetoresistance (AMR) เป็นเซนเซอร์วัดค่าความเข้มสนามแม่เหล็กชนิดหนึ่ง มีกระบวนการผลิตที่ง่าย และมีจำหน่ายทั่วไป เมื่อเปรียบเทียบกับเซนเซอร์วัดค่าความเข้มสนามแม่เหล็กอื่น ๆ AMR มีความเสถียรมากที่สุด และมีขนาดเล็ก โดยส่วนใหญ่ถูกใช้กับระบบนำทาง ซึ่งมีราคาถูกกว่าเซนเซอร์ GMR

3.2 Giant Magneto Resistance (GMR) เป็นเซนเซอร์ที่ใช้วัสดุที่มีความต้านทานไฟฟ้าลดลงอย่างมาก ขณะอยู่ในสนามแม่เหล็ก ถูกค้นพบโดย Baibich และคณะ (1988)

โครงสร้างของ GMR ประกอบด้วย 4 ชั้นที่มีคุณสมบัติต่างกัน (Mohamadabadi, 2013) ดังภาพ 5



ภาพ 5 โครงสร้างของเซนเซอร์ GMR (Mohamadabadi, 2013)

จากภาพ 5 จุดสำคัญของ GMR คือความหนาของชั้น ซึ่งทำให้เกิดความต้านทานไฟฟ้าที่แตกต่างกัน โดยชั้นโครงสร้าง Antiferromagnetic ใช้เพื่อกำหนดทิศทาง ชั้นที่ไม่ใช่แม่เหล็กที่มีการนำไฟฟ้าสูงจะแยกชั้นแม่เหล็กอื่น ๆ ออกจากกัน เซนเซอร์ GMR เป็นเซนเซอร์ที่มีความไวสูง และมีความเสถียรของอุณหภูมิ แต่สามารถถูกทำลายได้ง่ายในสนามแม่เหล็กแรงสูง เนื่องจากมีขั้วเดียว และพลังงานที่สูญเสียจะย้อนกลับไปในรูปแบบของความร้อน (Mohamadabadi, 2013)

4. Inertial Measurement Unit (IMU) เป็นหน่วยตรวจวัดความเคลื่อนไหวภายใน มีหน้าที่รายงานตำแหน่งการขยับเขยื้อนของวัตถุตามแนวการเคลื่อนไหวทั้งทิศทางแกน X Y และ Z รวมถึงอัตราการเร่งตามแกนต่าง ๆ ซึ่งประกอบด้วย 2 ส่วนหลัก คือ Accelerometer และ Gyroscope 3 ทิศทาง โดยมีขนาดเล็กเท่า Gyroscope ตัวอย่างเช่น อุปกรณ์ BMI263 ผลิตโดย Bosch SensorTec, Cornering ABS, ระบบ Traction Control ฯลฯ

### ชุดข้อมูลสาธารณะสำหรับกิจกรรมของมนุษย์

ชุดข้อมูลสาธารณะ เป็นชุดข้อมูลขนาดใหญ่ที่มีการจัดข้อมูลเป็นชุด ซึ่งมีลักษณะโครงสร้างของข้อมูลที่ถูกต้อง แต่ละชุดข้อมูลจะประกอบด้วยข้อมูล และรายละเอียดต่าง ๆ ที่เกี่ยวกับชุดข้อมูลนั้น ๆ เพื่อนำไปใช้ในการวิเคราะห์ และประมวลผลได้ โดยชุดข้อมูลสาธารณะนำมาใช้กันอย่างแพร่หลายทางด้านการเรียนรู้ของเครื่อง (Machine learning) และการเรียนรู้เชิงลึก (Deep learning) เป็นชุดข้อมูลสาธารณะสำหรับกิจกรรมของมนุษย์ส่วนใหญ่ ซึ่งถูกเก็บไว้ที่ศูนย์การเรียนรู้ของเครื่องและระบบอัจฉริยะที่มหาวิทยาลัยแคลิฟอร์เนีย เออร์ไวน์ ตัวอย่างชุดข้อมูลมีดังนี้

1. The UniMiB SHAR Datset (UniMiB SHAR) ชุดข้อมูลสาธารณะที่มาจากเซนเซอร์วัดความเร่งบนสมาร์ตโฟน โดยมีตัวอย่างของข้อมูลถึง 11771 ตัวอย่าง จาก 30 กิจกรรม โดยกลุ่มตัวอย่างมีอายุระหว่าง 18 ถึง 60 ปี กลุ่มตัวอย่างจะถูกแบ่งเป็น 17 ประเภท ประกอบด้วยกิจกรรมในชีวิตประจำวัน 9 ประเภท และกิจกรรมอื่น ๆ อีก 8 ประเภท (Micucci Daniela , Mobilio Marco และNapoletano Paolo, 2017)

2. The UCI Dataset (UCI) สร้างมาจากการบันทึกของ 30 กิจกรรมในชีวิตประจำวัน กลุ่มอาสาสมัครจำนวน 30 คน อยู่ในช่วงอายุ 19 ถึง 48 ปี ประกอบด้วยกิจกรรม 6 ประเภท ได้แก่ เดิน เดินขึ้นบันได เดินลงบันได นั่ง ยืน และนอน โดยอาสาสมัครทำการสวมสมาร์ตโฟนซัมซุงกาแล็กซี่เอสทู (Samsung Galaxy S II) ที่เอว เก็บข้อมูลจากเซนเซอร์วัดความเร่ง และเซนเซอร์วัดการหมุนที่ฝังไว้กับอุปกรณ์ โดยจับภาพความเร่ง 3 แกน และความเร็ว 3 แกน ในอัตราคงที่ 50 Hz ในการทดลองได้มีการบันทึกวิดีโอเพื่อติดป้ายกำกับข้อมูล ชุดข้อมูลนี้ถูกสุ่มแบ่งเป็น 2 ชุด โดย 70% ของข้อมูลมาจากอาสาสมัคร และ 30% ของข้อมูลสำหรับการทดสอบ (Anguita Davide และคณะ, 2012)

3. The OPPORTUNITY Dataset (OPPORTUNITY) ชุดข้อมูลสาธารณะสำหรับกิจกรรมของมนุษย์จากเซนเซอร์สวมใส่ ซึ่งมาจากงานวิจัยเชิงเทคนิคการเรียนรู้ของเครื่องสำหรับการรู้จำกิจกรรมของมนุษย์ โดยข้อมูลมาจากเซนเซอร์สวมใส่ IMU 7 ตัว เซนเซอร์ Accelerometer 3 ทิศทาง 12 ตัว และข้อมูลองค์ความรู้ท้องถิ่น 3 ทิศทาง 4 ตัว เซนเซอร์ของวัตถุ 12 ตัว เซนเซอร์รอบข้างจาก สวิตช์ 13 ตัว และเซนเซอร์ Accelerometer 3 ทิศทาง บันทึกโดยอาสาสมัคร 4 คน โดยมีการบันทึก 6 รอบ ใน 5 รอบเป็นกิจกรรมทั่วไปในชีวิตประจำวัน และอีก 1 รอบเป็นกิจกรรมตามที่กำหนด กิจกรรมในสถานการณ์ต่าง ๆ แบ่งเป็น 3 ระดับ ได้แก่

ระดับต่ำ 13 กิจกรรม 23 วัตถุ ระดับกลาง 17 กิจกรรม และระดับสูง 5 กิจกรรม กิจกรรมทั่วไปในชีวิตประจำวันประกอบด้วย การนอน การยืน การเดิน การเตรียมกาแฟ การจับกาแฟ การดื่มกาแฟ การเตรียมแซนวิช การทำความสะอาด เป็นต้น กิจกรรมตามที่กำหนดประกอบด้วย การเปิด-ปิดตู้เย็น การเปิด-ปิดตู้เย็น การเปิด-ปิดลิ้นชักที่ความสูงต่างกัน การเปิด-ปิดประตู การเปิด-ปิดไฟ การทำความสะอาดโต๊ะ การดื่มขณะยืน การดื่มขณะนั่ง เป็นต้น (Roggen Daniel และคณะ, 2010)

4. The PAMAP2 Dataset (PAMAP2) ประกอบด้วยกิจกรรมทางกายภาพ 18 แบบ ดำเนินการโดยอาสาสมัคร 9 คน ที่สวมเซ็นเซอร์ IMU และเครื่องวัดอัตราการเต้นของหัวใจในการสวมตัวอย่าง 100 Hz โดยตำแหน่งของเซ็นเซอร์มี 3 ที่ เหนือข้อมือข้างที่ถนัด หน้าอก และข้อเท้าข้างที่ถนัด เครื่องมือใช้วัดอัตราการเต้นของหัวใจมีความถี่ในการสวมตัวอย่าง ประมาณ 9 Hz มีกิจกรรมดังนี้ 1 นอน นั่ง ยืน เดิน วิ่ง ปั่นจักรยานเดินนอร์ดิก ดูทีวี คอมพิวเตอร์ทำงาน ขับรถ ขึ้นบันได บันได เครื่องดูดฝุ่น รีดผ้า ซักผ้า พับ ทำความสะอาดบ้าน เล่นฟุตบอล กระโดดเชือก และกิจกรรมอื่น ๆ (กิจกรรมชั่วคราว) (Reiss Attila และ Stricker Didier, 2012)

5. The WISDM Dataset (WISDM) เป็นชุดข้อมูลอนุกรมเวลาที่มาจากเซ็นเซอร์วัดความเร่ง และเซ็นเซอร์วัดการหมุน เก็บข้อมูลมาจากสมาร์ทโฟน (Google Nexus 5/5x) และสมาร์ทวอตช์ (LG G Watch) ที่อัตรา 20 Hz โดยอาสาสมัคร 51 คน 18 กิจกรรม ได้แก่ เดิน วิ่ง ขึ้นบันได นั่ง ยืน การพิมพ์ แปรงฟัน กินชุป กินมันฝรั่ง กินพาสต้า ดื่มจากถ้วย กินแซนวิช ตะโพกฟุตบอล เล่นเทนนิส เลี้ยงลูกบอล การเขียน ปรบมือ และพับผ้า กิจกรรมละ 3 นาที (Weiss Gary M, Yoneda Kenichi และ Hayajneh Thair, 2019)

6. The Daphnet freezing of gait Dataset (Daphnet) ชุดข้อมูลที่มาจกเซ็นเซอร์ Accelerometer 3 ตัวติดอยู่ที่สะโพก และขาของผู้ป่วยพาร์กินสันที่มีปัญหาการเดินติดขัด (freezing of gait : FoG) ในชุดข้อมูลจะบันทึกการเดินเป็นเส้นตรง เดินสลับไปมา และกิจกรรมในชีวิตประจำวัน ชุดข้อมูลนี้เป็นผลที่มาจากการทำงานร่วมกันระหว่างห้องปฏิบัติการฟื้นฟูผู้ป่วยที่มีปัญหาการเดิน และกายภาพบำบัด ศูนย์การแพทย์เทลฮาวิฟ ไชราสกี ประเทศอิสราเอล และห้องปฏิบัติการอุปกรณ์ไอทีแบบสวมใส่ สถาบันเทคโนโลยีแห่งสหพันธ์สวิสในซูริค ถูกบันทึกโดยศูนย์การแพทย์เทลฮาวิฟ ไชราสกีในปี 2008 (Bächlin Marc และคณะ, 2010)

7. The Wetlab Dataset (Wetlab) ชุดข้อมูลที่ประกอบด้วยกิจกรรมที่แตกต่างกัน ดังนี้ การตัด การพลิกกลับด้าน การปกปิดเปลือก การบิด การวัดปริมาตรของสารละลาย การเท การกวน การย้าย และการกระทำที่ไม่สามารถระบุได้ จากอาสาสมัคร 22 คน สวมเซนเซอร์ไว้ที่ข้อมือ ข้อมูลมาจากเซนเซอร์ accelerometer 3 ทิศทาง สุ่มตัวอย่างที่ 50 Hz รวม 1258 นาที ของข้อมูลจากเซนเซอร์ (Scholl Philipp M, Wille Matthias และ Van Laerhoven Kristof, 2015)

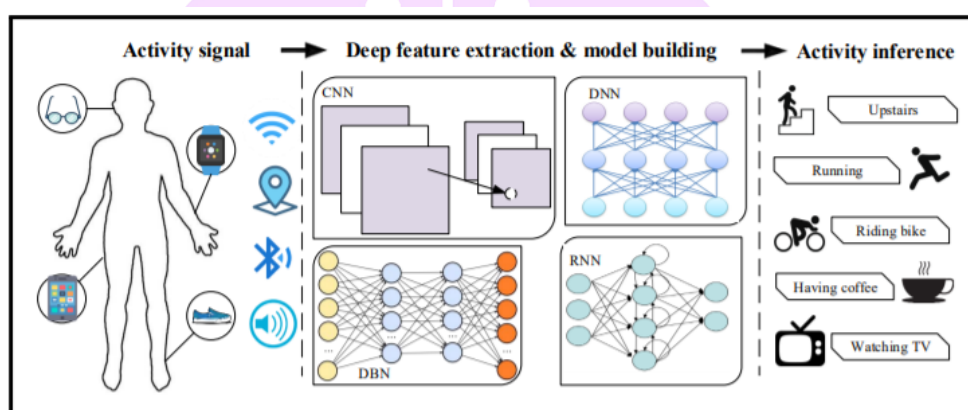
8. The RealWorld HAR Dataset (RWHar) มาจากเซนเซอร์ Accelerometer ประกอบด้วยกิจกรรม 8 กิจกรรม ได้แก่ เดินขึ้นบันได เดินลงบันได กระโดด นอน ยืน นั่ง วิ่ง และเดิน และกิจกรรมที่ระบุค่าไม่ได้ จากอาสาสมัคร 15 คน (Sztyler Timo และ Stuckenschmidt Heiner, 2016)

9. The Heterogeneity Activity Recognition Dataset (HHAR) ชุดข้อมูลที่มาจากสมาร์ทโฟน สมาร์ทวอตช์ และแท็บเล็ต จากเซนเซอร์ Accelerometer โดยตรวจจับ และแยกกิจกรรมของผู้ใช้ 6 กิจกรรม การปั่นจักรยาน การนั่ง การยืน การเดิน การขึ้นบันได และการลงบันได จากอาสาสมัคร 9 คน โดยใช้สมาร์ทโฟนรวมทั้งหมด 31 เครื่อง สมาร์ทวอตช์ 4 เครื่อง และแท็บเล็ต 1 เครื่อง โดย 8 เครื่อง ถูกเก็บไว้ในกระเป๋า และสวมสมาร์ทวอตช์ สองเรือนที่แขนแต่ละข้าง ผู้เข้าร่วมดำเนินกิจกรรม กิจกรรมละ 5 นาที สุ่มตัวอย่างที่ 200 Hz สำหรับ LG Nexus 4, 150 Hz สำหรับ Samsung Galaxy S3, 100 Hz สำหรับ Samsung Galaxy S3, 50 Hz สำหรับ Samsung Galaxy S plus (Stisen Allan และคณะ, 2015)

### การเรียนรู้เชิงลึก

การเรียนรู้เชิงลึก (Deep Learning) ออกแบบมาจากการทำงานของสมองมนุษย์ โดยเป็นการฝึกฝนคอมพิวเตอร์ให้สามารถทำงานได้เหมือนมนุษย์ จำลองมาจากเซลล์สมองของมนุษย์หลาย ๆ เซลล์รวมเป็นโครงข่ายใหญ่ ๆ ที่ซับซ้อน การเรียนรู้เชิงลึกเป็นแบบจำลองที่ทำให้เครื่องจักรสามารถตัดสินใจได้เอง โดยมาจากการเรียนรู้ของเครื่องที่นำความรู้ทางด้านสถิติมาประยุกต์ใช้ในการวิเคราะห์ข้อมูล และสร้างแบบจำลองขึ้นมาเพื่อทำนายผลลัพธ์จากข้อมูล (ภูมिरพี ภูมิดำ, 2019) การเรียนรู้เชิงลึกเป็นโครงข่ายประสาทเทียม (Neural Network) ที่เลียนแบบการทำงานของสมองมนุษย์ ซึ่งมีการทำงานที่ซับซ้อน และสามารถวิเคราะห์สิ่งต่าง ๆ ได้อย่างมีประสิทธิภาพ โดยโครงข่ายประสาทเทียมทำการจำลองการทำงานของเซลล์ประสาทขึ้นมา โดยมีจำนวนเซลล์หลายเซลล์ แต่ละเซลล์จะเชื่อมต่อกันเพื่อส่งข้อมูลไปหากัน สำหรับใช้ในการตัดสินใจ (Hecht-Nielsen Robert, 1992)

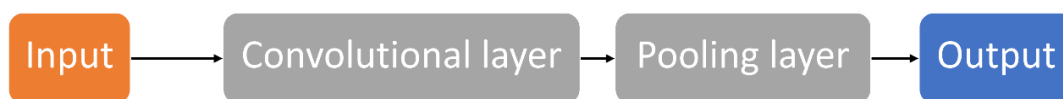
แบบจำลองที่ได้รับความนิยม ได้แก่ Deep Neural Network, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM) และ Gated Recurrent Units (GRU) เป็นต้น การเรียนรู้เชิงลึกได้รับความนิยมเพิ่มมากขึ้น เนื่องจากมีประสิทธิภาพในการวิเคราะห์สูง และมีการพัฒนาเทคโนโลยีทางด้านฮาร์ดแวร์มากขึ้น นอกจากนี้ยังมีการนำไปประยุกต์ใช้ใน Graphic Processing Unit (GPU) ในการเรียนรู้เชิงลึก ซึ่งประมวลผลได้เร็วกว่า Central Processing Unit (CPU) (Dean Jeffrey และคณะ, 2012)



ภาพ 6 การรู้จำกิจกรรมของมนุษย์จากเซนเซอร์โดยใช้การเรียนรู้เชิงลึก  
(Wang และคณะ, 2019)

Hinton Geoffrey E, Osindero Simon และTeh Yee-Whye (2006) การเรียนรู้เชิงลึกสามารถเรียนรู้คุณลักษณะต่าง ๆ ได้โดยอัตโนมัติผ่านโครงข่าย และสามารถแยกคุณลักษณะที่มีความลึก ซึ่งเหมาะกับการรู้จำกิจกรรมที่ซับซ้อนมากขึ้น เมื่อต้องใช้ข้อมูลจำนวนมากที่ไม่มีการระบุคำตอบมาให้

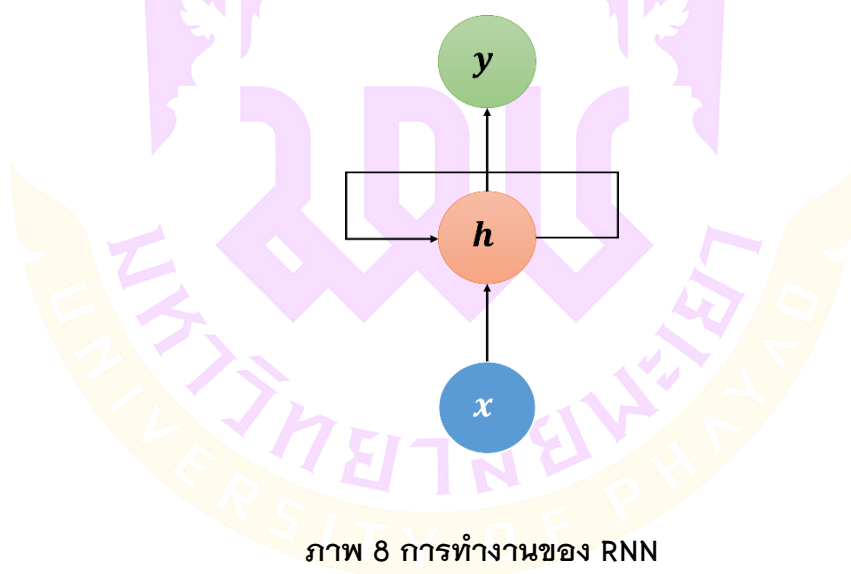
1) โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution neural Network : CNN) เป็นโครงข่ายประสาทเทียมแบบพิเศษ ซึ่งสามารถจำแนกข้อมูลประเภทรูปภาพได้ดี เนื่องจากแบบจำลองโครงข่ายประสาทเทียมนี้จำลองการมองเห็นคล้ายมนุษย์ กล่าวคือสกัดเอาคุณลักษณะสำคัญต่าง ๆ เพื่อเรียนรู้ลักษณะของภาพนั้น ๆ ได้อย่างแม่นยำ โดยโครงสร้างของแบบจำลองนี้จะถูกแบ่งเป็นชั้น ๆ แสดงดังภาพ



ภาพ 7 การทำงานของ CNN

จากภาพอธิบายภาพรวมการทำงานของแบบจำลอง เริ่มจากรับข้อมูลนำเข้า หรือข้อมูล Input ชั้นถัดมาเป็นชั้นแรกเรียกว่า Convolution layer

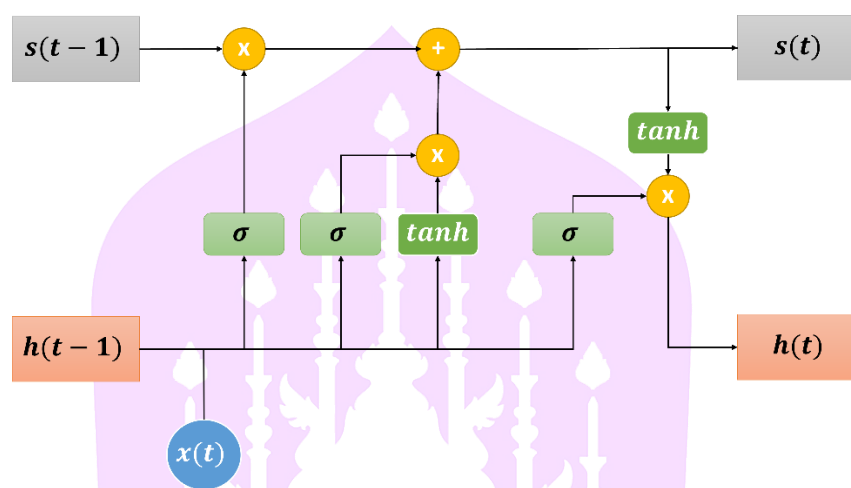
2) โครงข่ายประสาทเทียมแบบวนกลับ (Recurrent Neural Network : RNN) เป็นแบบจำลองที่นำมาใช้ในการเรียนรู้จำเลียง และการประมวลผลทางด้านภาษธรรมชาติ โดยกระบวนการทำงานของแบบจำลองจะนำเอาผลลัพธ์ที่ได้จากการคำนวณของข้อมูลขาเข้ามาใช้ซ้ำหรือเรียกว่าหลักการ Feed ซึ่งวิธีการนี้มีประโยชน์กับข้อมูลที่มีความต่อเนื่อง หรือข้อมูลที่เปลี่ยนแปลงไปตามเวลา (Time Series) ลักษณะการทำงานของแบบจำลองแสดงดังภาพ



ภาพ 8 การทำงานของ RNN

ข้อมูลขาเข้าที่ถูกรับเข้ามา (x) จะถูกคำนวณและกลับมาเป็น input ใหม่ที่ตรงกับ input ปกติ ซึ่งเรียกว่า Hidden State (h) การทำงานจะเป็นการวนซ้ำ โดยที่มีการใช้ข้อมูล hidden state ก่อน ๆ มา และวนลูปซ้ำไปเรื่อย ๆ

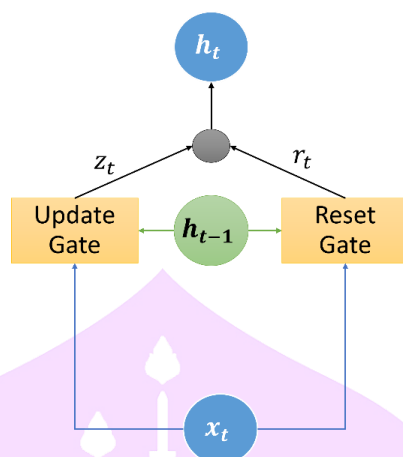
3) Long Short-Term Memory (LSTM) แบบจำลองนี้เป็นแบบจำลองที่พัฒนามาเพื่อแก้ไขปัญหา Vanish Gradient ซึ่งปัญหานี้ถูกพบเมื่อใช้แบบจำลอง RNN กับข้อมูลที่มีลักษณะที่ต่อเนื่องกันยาว ๆ เช่นเสียงที่มีลักษณะยาว ๆ ในแบบจำลองนี้มีลักษณะการทำงานดังภาพ



ภาพ 9 การทำงานของ LSTM

จากภาพ แบบจำลอง LSTM จะประกอบด้วย 3 ชั้น ได้แก่ input gate, output gate และ forget gate ในชั้นแรกเริ่มรับข้อมูลเข้ามาในชั้น input จากนั้นในชั้น output gate จะทำการตัดสินใจว่าค่าที่รับมานั้นจะเก็บไว้หรือจะทำการวนซ้ำกลับไปใน LSTM หากแบบจำลองตัดสินใจว่าจะวนซ้ำข้อมูลจะถูกวนกลับมาเพื่อเข้าสู่ชั้น forget gate ชั้นนี้จะทำการตัดสินใจว่าค่าที่รับมานั้นจะลบค่านั้นออกไปหรือเก็บค่านั้นไว้ หากค่านั้นถูกตัดสินใจว่าเก็บไว้ จะถูกรอเพื่ออัปเดตจาก input gate และ input gate จะตัดสินใจว่าจะอัปเดตค่านั้นด้วยค่าอะไร และค่านั้นจะถูกส่งไปที่ output gate เพื่อตัดสินใจอีกครั้ง โดยจะวนซ้ำไปอีกรอบเรื่อย ๆ

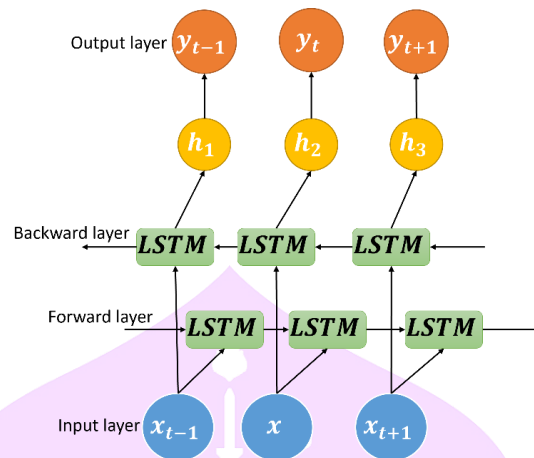
4) หน่วยการเวียนกลับแบบมีประตู (Gate Recurrent Unit : GRU) แบบจำลองนี้ถูกพัฒนาต่อยอดมาจาก LSTM ซึ่งมีหลักการคล้ายกับ LSTM แต่จะมีรวดเร็วกว่า เนื่องจากมีการลดความซับซ้อน และการตัดกระบวนการทำงานของ input gate และ output gate ออกจากการวนซ้ำ และเปลี่ยนมาใช้วิธีการ reset gate และ update gate (แทน ภูมिरพี ,2562) แบบจำลองนี้มีลักษณะการทำงานดังภาพ



ภาพ 10 การทำงานของ GRU

จากภาพ แบบจำลอง GRU จะประกอบด้วย 3 ชั้น ได้แก่ input gate, output gate และ forget gate ในชั้นแรกเริ่มรับข้อมูลเข้ามาในชั้น input จากนั้นจะตัดสินใจว่าข้อมูลที่รับเข้ามา นั้นจะนำไปวนซ้ำหรือไม่ หากวนซ้ำข้อมูลจะถูกส่งไป reset gate ชั้นนี้จะตัดสินใจว่าข้อมูล โหนดจะถูกลบ หรือถูกเก็บไว้ หลังจากนั้นจะไปที่ update gate เพื่อตัดสินใจว่าจะอัปเดตข้อมูล โหนด และสุดท้ายจะตัดสินใจว่าจะวนซ้ำหรือไม่

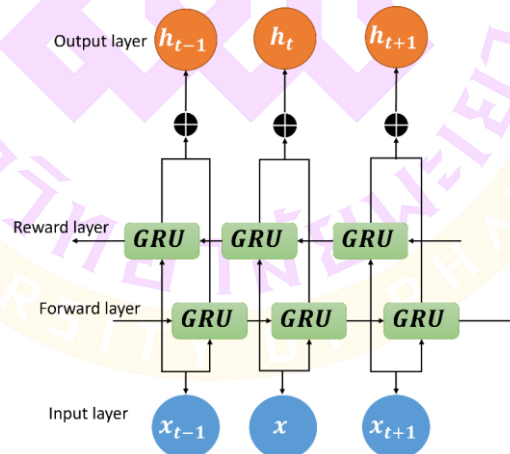
5) Bi-directional LSTM (Bi-LSTM) มีหลักการทำงานเหมือนกับ LSTM แต่เพิ่มการ ป้อนข้อมูลย้อนกลับเข้าไป ซึ่งจะทำให้ขั้นตอนในการทำงานช้ากว่า LSTM ลักษณะการทำงาน ของแบบจำลองแสดงดังภาพ



ภาพ 11 การทำงานของ Bi-LSTM

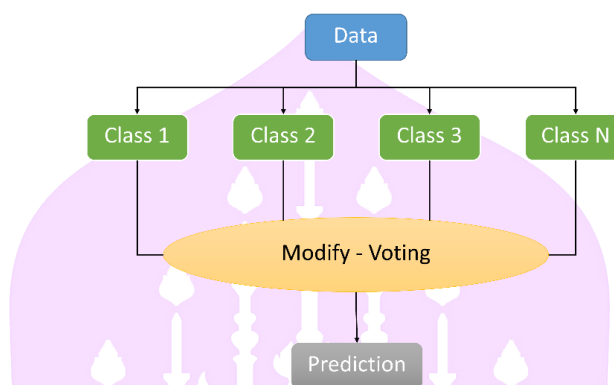
จากภาพ แบบจำลอง Bi-LSTM เป็น LSTM แบบสองทาง ซึ่งนำ LSTM มาต่อกันและส่งค่าแบบสองทางทั้ง Forward และ Backward

6) Bi-directional GRU (Bi-GRU) มีหลักการการทำงานเหมือนกับ GRU แต่เพิ่มการป้อนข้อมูลย้อนกลับเข้าไป ซึ่งจะทำให้ขั้นตอนในการทำงานช้ากว่า GRU ลักษณะการทำงานของแบบจำลองแสดงดังภาพ



ภาพ 12 การทำงานของ Bi-GRU

7) Ensemble model เป็นการใช้แบบจำลองการจัดหมวดหมู่หลาย ๆ แบบจำลอง โดยใช้เทคนิคหลัก ๆ 3 เทคนิคดังนี้ Voting Ensemble, Bootstrap Aggregating (Bagging) และ Random Forest



ภาพ 13 การทำงานของ Random Forest

#### มาตรวัดประสิทธิภาพของแบบจำลองแบบหมวดหมู่

การที่จะนำแบบจำลองไปใช้งานได้นั้น จำเป็นต้องมีการวัดประสิทธิภาพของแบบจำลองก่อน เพื่อดูว่าแบบจำลองนั้นมีประสิทธิภาพและประสิทธิผลเพียงพอที่จะนำมาพัฒนา หรือใช้งานในด้านอื่น ๆ ต่อไป ซึ่งในการวัดประสิทธิภาพนั้นได้มีมาตรวัดของแต่ละประเภท ดังนี้

1) Confusion Metrix เป็นตารางในการวัดความสามารถของแต่ละคลาส หรือเรียกว่า เมทริกซ์ความสัมพันธ์ของการประเมินคุณภาพแต่ละคลาส โดยที่การแสดงผลในแนวทแยงแสดงคำตอบที่ถูกต้อง นอกแนวทแยงเป็นคำตอบที่ไม่ถูกต้อง

2) Precision เป็นค่าความแม่นยำที่มาจากการนำค่าของข้อมูลที่ทำนายแล้วถูกต้อง เมื่อเทียบกับคำตอบเทียบกับค่าของข้อมูลที่อยู่ในคำตอบแต่ไม่มีในการทำนาย มีสูตรการคำนวณดังนี้

$$Precision = TP / (TP + FP) \quad (2)$$

3) Recall เป็นค่าความถูกต้องที่มาจากการนำค่าของข้อมูลที่ทำนายแล้วถูกต้องเมื่อเทียบกับคำตอบมาเทียบกับค่าของข้อมูลที่ทำนายแล้วไม่ถูกต้องเมื่อเทียบกับคำตอบ มีสูตรการคำนวณดังนี้

$$Recall = TP / (TP + FN) \quad (3)$$

3) Accuracy เป็นค่าความแม่นยำของแบบจำลอง มีสูตรการคำนวณดังนี้

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (4)$$

4) F-Measure คือค่าเฉลี่ยของ Precision และ Recall

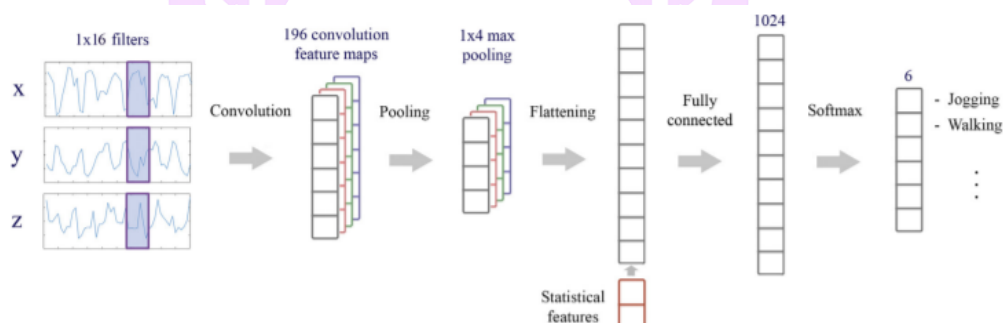


## งานวิจัยที่เกี่ยวข้อง

จากการศึกษางานวิจัยที่เกี่ยวข้องกับการเรียนรู้เชิงลึกในการรู้จำกิจกรรมของมนุษย์ ได้มีวิธีการต่าง ๆ อย่างหลากหลายที่นำมาแก้ปัญหาการรู้จำกิจกรรมของมนุษย์ ซึ่งผู้วิจัยได้สรุปเป็น 3 วิธีการ ดังนี้

### 1. Real-time human activity recognition from accelerometer data using Convolutional Neural Networks

จากงานวิจัย (Ignatov Andrey, 2018) ได้นำเสนอการรู้จำกิจกรรมของมนุษย์แบบเรียลไทม์จากข้อมูลเซนเซอร์วัดความเร่งโดยใช้โครงข่ายประสาทเทียม งานวิจัยนี้ใช้ชุดข้อมูลสาธารณะ WISDM และ UCI ที่มาจากอาสาสมัครจำนวน 36 และ 30 คนตามลำดับ โดยมีกิจกรรมทั้งหมด 6 กิจกรรมในชุดข้อมูล วิธีการในงานนี้ใช้โครงข่ายประสาทเทียม โดยทดลองปรับค่าพารามิเตอร์ต่าง ๆ เช่น จำนวน และขนาดของตัวกรองโครงข่ายประสาทเทียม, ชั้นกลางของโครงข่าย และลดค่าบางค่าลง ดังภาพ 14

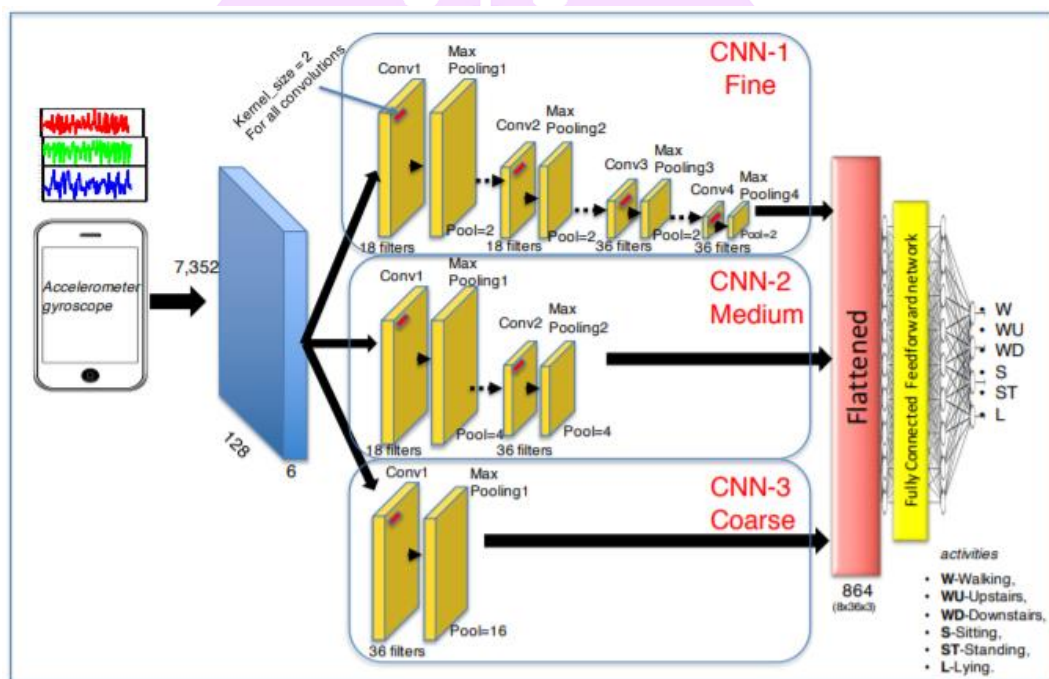


ภาพ 14 สถาปัตยกรรมของแบบจำลอง CNN (Ignatov Andrey, 2018)

ผลการทดลองกับชุดข้อมูล 2 ชุด ประสิทธิภาพของแบบจำลองมีความแม่นยำอยู่ที่ 97.63% และ 90.42% จากชุดข้อมูล UCI และ WISDM ตามลำดับ แบบจำลองที่ได้นำเสนอนั้นมีขนาดของแบบจำลองเล็ก เพื่อให้สามารถทำงานบนอุปกรณ์ขนาดเล็กแบบเรียลไทม์ได้ และมีความแม่นยำสูง แต่สามารถใช้ในการรู้จำกิจกรรม ที่มีขนาดของอนุกรมเวลาสูงสุดไม่เกิน 1 วินาที หากเวลามากกว่า 1 วินาที อาจส่งผลต่อความแม่นยำในการรู้จำกิจกรรม

## 2. Coarse-Fine Convolutional Deep-Learning Strategy for Human Activity Recognition

จากงานวิจัย (Avilés-Cruz Carlos และคณะ, 2019) ได้นำเสนอการเรียนรู้เชิงลึกโดยใช้กลยุทธ์ของโครงข่ายประสาทเทียม ชุดข้อมูลสาธารณะ UCI ได้ถูกนำมาใช้ทดสอบในงานวิจัยนี้ ในชุดข้อมูลที่กล่าวมานั้น มาจากเซนเซอร์วัดความเร่ง และเซนเซอร์วัดการหมุนแบบจำลองที่เสนอนั้นมีหลักการทำงานดังภาพ 15



ภาพ 15 การทำงานของโครงข่ายประสาทเทียม Coarse-fine Carlos (Avilés-Cruz Carlos และคณะ, 2019)

งานวิจัยนี้ได้เสนอแบบจำลองที่มีการรวมกันของ 3 โครงข่ายประสาทเทียม ได้แก่ fine-Fine-CNN, medium-CNN และ Coarse-CNN โดยใช้ฟังก์ชันการสูญเสีย (loss entropy function) ดังสมการ

$$\begin{aligned}
 \mathcal{T}(\theta_1, \theta_2, \theta_3, W) &= - \sum_{i=1}^3 \log[\hat{p}(m_h/\theta_i)] - \log[\hat{p}(m_h/W)] \\
 &= - \log[\hat{p}(m_h/\theta_1)] - \log[\hat{p}(m_h/\theta_2)] - \\
 &\quad \log[\hat{p}(m_h/\theta_3)] - \log[\hat{p}(m_h/W)]
 \end{aligned} \tag{5}$$

โดยที่  $W$  คือ พารามิเตอร์ของชั้นสุดท้าย

$m_h$  คือ ประเภทการเคลื่อนไหว

$\hat{p}$  คือ พังก์ชันความน่าจะเป็นที่มีเงื่อนไขสำหรับประเภทการเคลื่อนไหวที่

กำหนดเป็นพารามิเตอร์  $W$

$\theta_1$  คือพารามิเตอร์ของ fine-CNN

$\theta_2$  คือพารามิเตอร์ของ medium-CNN

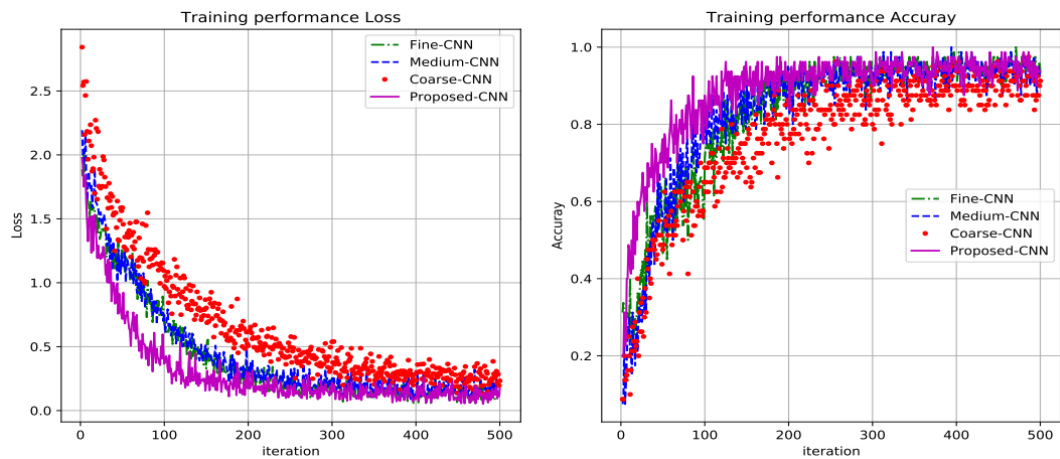
$\theta_3$  คือพารามิเตอร์ของ coarse-CNN

จากชุดข้อมูล UCI มาจากอาสาสมัครจำนวน 30 คน มีอายุระหว่าง 19–49 ปี โดยนำสมาร์ตโฟน (Samsung Galaxy SLII) จัดตำแหน่งแนวตั้งใส่ไว้ในกระเป๋า และอาสาสมัครทดลองทำกิจกรรม 6 กิจกรรม ดังนี้ การเดิน การเดินขึ้นบันได การเดินลงบันได การนั่ง การยืน และการนอน ข้อมูลจะถูกเก็บจากเซนเซอร์แบบ 3 แกนจากเซนเซอร์วัดความเร็ว และวัดการหมุน โดยสุ่มตัวอย่างอัตราคงที่ 50 Hz 1 กิจกรรมจะถูกแบ่งออกเป็นหน้าต่างละ 2.56 วินาที จะทำให้ได้ทั้งหมด 128 ตัวอย่าง ( $2.56s * 50Hz = 128$ ) ข้อมูลจะถูกแบ่งเป็น 2 ชุด ชุดแรกเป็นชุดที่นำมาฝึกฝน จำนวน 70% ประกอบด้วยชุดข้อมูลจากอาสาสมัครจำนวน 21 คน และชุดที่สองเป็นชุดที่นำมาทดสอบ ประกอบด้วยชุดข้อมูลจากอาสาสมัครจำนวน 9 คน แสดงดังตาราง 1

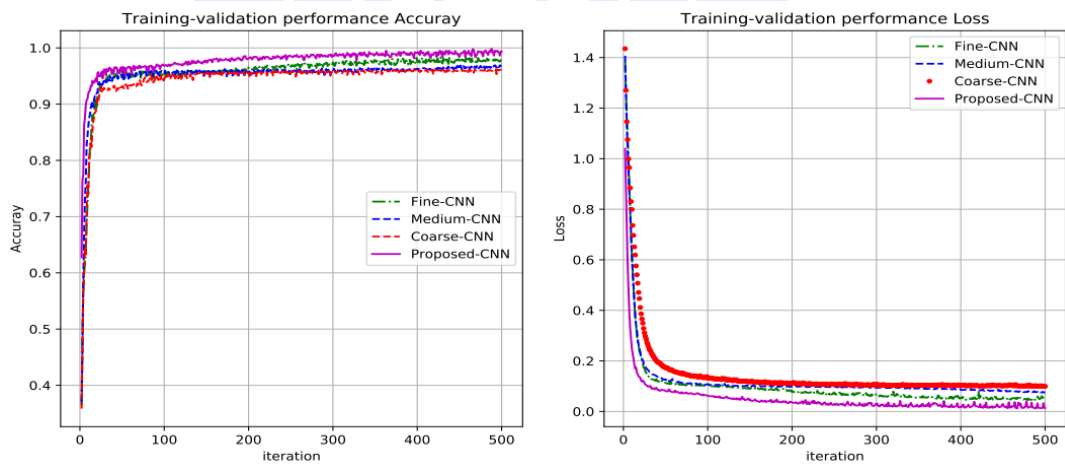
ตาราง 1 การแบ่งชุดข้อมูลการฝึกอบรมและทดสอบของชุดข้อมูล UCI

	Training	Testing
walking	1226	496
walking-upstairs	1073	471
walking-downstairs	986	420
sitting	1286	491
standing	1374	532
laying	1407	537
Total	7352	2947

งานวิจัยนี้ได้ทำการทดสอบโดยใช้แบบจำลอง 4 แบบจำลอง ดังนี้ Fine-CNN, medium-CNN, Coarse-CNN และ Coarse- Fine CNN โดยค่าความสูญเสีย และความแม่นยำ แสดงดังภาพ 16 และ 17



ภาพ 16 ค่าความสูญเสียของชุดฝึกฝน Coarse-Fine



ภาพ 17 ค่าความแม่นยำของชุดทดสอบ Coarse-Fine

จากภาพแสดงให้เห็นว่าแบบจำลองที่งานวิจัยนี้นำเสนอในชุดทดสอบมีความแม่นยำที่สูง และมีความสูญเสียที่ต่ำ ผลการทดสอบประสิทธิภาพการจัดหมวดหมู่กิจกรรมจำนวน 6 กิจกรรม มีความประสิทธิภาพเฉลี่ย 100% ในทุกกิจกรรม แสดงดังตาราง 2

ตาราง 2 ประสิทธิภาพการทดสอบของ 6 กิจกรรม

	Walking	Ascending Stairs	Descending Stairs	Sitting	Standing	Laying
walking	100	0	0	0	0	0
walking-upstairs	0	100	0	0	0	0
walking-downstairs	0	0	100	0	0	0
sitting	0	0	0	100	0	0
standing	0	0	0	0	100	0
laying	0	0	0	0	0	100

ประสิทธิภาพของแบบจำลองให้ประสิทธิภาพที่สูงกว่าแบบจำลองอื่น ๆ แต่ในชุดข้อมูลกิจกรรมของมนุษย์ที่มีความซับซ้อนมาก ยังไม่สามารถจัดหมวดหมู่ได้ดี

### 3. A Lightweight Deep Learning Model for Human Activity Recognition on Edge Devices

Agarwal Preeti Alam Mansaf (2020) ได้นำเสนอแบบจำลองการเรียนรู้เชิงลึกที่มีน้ำหนักเบาสำหรับการรู้จำกิจกรรมของมนุษย์บนอุปกรณ์ขนาดเล็ก งานวิจัยนี้ใช้ชุดข้อมูลสาธารณะ WISDM ที่มาจากเซนเซอร์วัดความเร่ง ที่มี 6 กิจกรรม โดยแบ่งข้อมูลเป็น 70 ต่อ 30 ของข้อมูลทั้งหมด สำหรับฝึกฝนและทดสอบ

โครงสร้างของแบบจำลองนี้มีขั้นตอน 5 ขั้นตอนดังนี้ การเก็บรวบรวมข้อมูล การเตรียมข้อมูล การสกัดคุณลักษณะ การจำแนกประเภทของกิจกรรม และประเมินผลของแบบจำลอง แสดงดังภาพ



ภาพ 18 ขั้นตอนพื้นฐานการเรียนรู้จำกิจกรรมของมนุษย์

ในส่วนของแบบจำลองที่เสนอคือ โครงข่ายประสาทเทียมแบบตื่นที่ออกแบบมาสำหรับแก้ปัญหาในข้อมูลที่เป็นลำดับชั้น (Recurrent neural network : RNN) ร่วมกับโครงข่ายประสาทเทียมความจำระยะสั้นยาว (Long short-term memory : LSTM) โดยมีโครงสร้างที่ตื่นเพียง 2 ชั้น และ 30 เซลล์ประสาททำให้สามารถติดตั้งบนอุปกรณ์ขนาดเล็กได้ ฟังก์ชันของ RNN แสดงดังสมการ

$$h_t = e(W_{hh}h_{t-1} + W_{ih}X_t + b_h) \quad (6)$$

$$y_t = e(W_{ho}h_t + b_o) \quad (7)$$

โดยที่  $X_t$  คือสถานะของข้อมูลขาเข้าปัจจุบัน

$h_{t-1}$  คือ hidden state ก่อนหน้า

$O_t$  คือค่า hidden ปัจจุบัน

$e$  คือ activation function

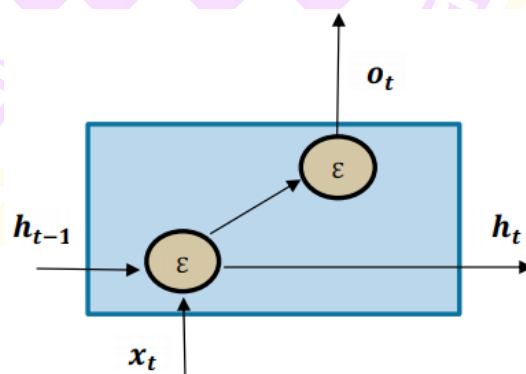
$W_{hh}$  คือ weight จาก hidden ไปยัง hidden state

$W_{ih}$  คือ weight จากข้อมูลขาเข้าไปยัง hidden state

$W_{ho}$  คือ weight จากสถานะซ่อนไปยัง hidden state

$b_h$  คือ ค่าคงที่ของ hidden state

$b_o$  คือค่าคงที่ของข้อมูลขาออก



ภาพ 19 โหนดของ RNN

จากปัญหา Vanishing Gradient Problem คือในระหว่างการเทรน Gradient มีขนาดเล็กลงเรื่อย ๆ จนเท่ากับ 0 ทำให้ค่า weight ไม่ถูกอัปเดตอีก ทำให้แบบจำลองฝึกฝนต่อไม่ได้ และ Exploding Gradient Problem คือในระหว่างการเทรน Gradient มีขนาดใหญ่ขึ้นเรื่อย ๆ จนมีค่าที่ไม่มีที่สิ้นสุด หรือค่าอนันต์ (infinity) ทำให้ระบบปรับค่าไม่ไหว และเทรนต่อไม่ได้ ซึ่ง LSTM มาแก้ไขปัญหานี้ LSTM จะประกอบด้วยพารามิเตอร์ และเกต เพื่อควบคุมพฤติกรรมของเซลล์หน่วยความจำ การคำนวณ hidden state ของ LSTM แสดงดังสมการ

$$i_t = \varepsilon_i(W_{xi}X_t + W_{hi}h_{t-1} + W_{ci}C_{t-1} + b_i) \quad (8)$$

$$f_t = \varepsilon_f(W_{xf}X_t + W_{hf}h_{t-1} + W_{cf}C_{t-1} + b_f) \quad (9)$$

$$c_t = f_t C_{t-1} + i_t \varepsilon_c(W_{xc}X_t + W_{hc}h_{t-1} + W_{cc}C_{t-1} + b_c) \quad (10)$$

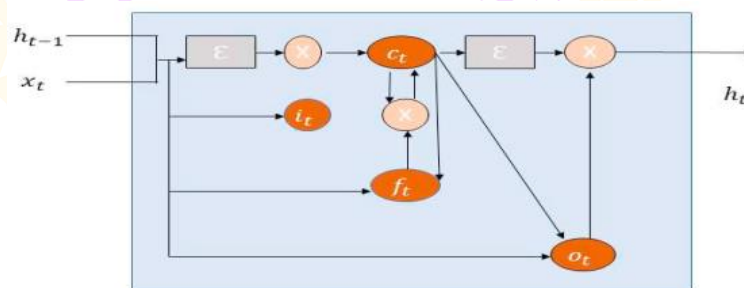
$$o_t = \varepsilon_o(W_{xo}X_t + W_{ho}h_{t-1} + W_{co}C_{t-1} + b_o) \quad (11)$$

$$h_t = o_t \varepsilon_h(c_t) \quad (12)$$

โดยที่  $W_{xi}, W_{hi}, W_{ci}, W_{xf}, W_{hf}, W_{cf}, W_{xc}, W_{hc}, W_{xo}$  และ  $W_{co}$  คือ weight

$b_i, b_o, b_c$  และ  $b_f$  คือ bias weight

การทำงานของ RNN-LSTM ค่าเช่นเซอร์วัดความเร่งจะถูกแบ่งเป็นหน้าต่าง T ข้อมูลขาเข้าจะแสดงเป็น  $(X_1, X_2, X_3, \dots, X_{T-1}, X_T)$  โดยที่  $X_t$  คือการอ่านค่า ณ เวลาใด ๆ จากนั้นจะรวมค่าโดยใช้ผลรวมของกฎ และข้อมูลขาออกจาก state ต่าง ๆ ในขั้นสุดท้าย เป็น  $O_T$  แสดงดังภาพ

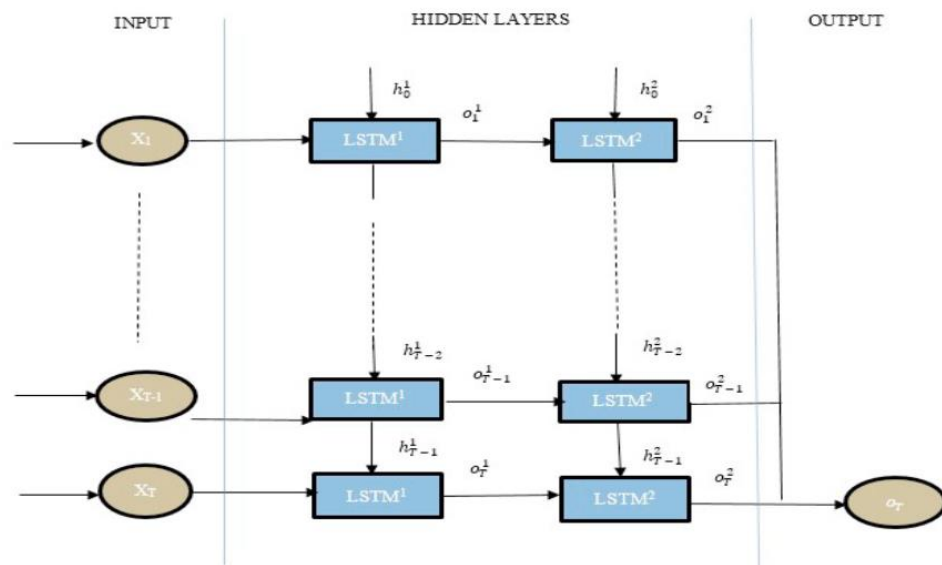


ภาพ 20 เซลล์หน่วยความจำของ LSTM

งานวิจัยนี้ได้นำเสนอแบบจำลอง LSTM ที่ประกอบด้วย hidden state 2 ชั้น 30 เซลล์ประสาท โดยมีข้อมูลขาเข้า ( $X_1, X_2, \dots, X_T$ ) ซึ่งเป็นสัญญาณที่เป็นเวลาแบบไม่ต่อเนื่องส่งเข้าไปชั้นที่ 1 ที่เวลา  $t$  โดยที่  $t = (1, 2, \dots, T)$  ผลลัพธ์ของชั้นแรกจะถูกสร้างสมการ

$$O_t^1, h_t^1, c_t^1 = LSTM^1(c_{t-1}^1, h_{t-1}^1, X_t, \theta^1) \quad (13)$$

โดยที่  $\theta^1$  คือพารามิเตอร์ของเซลล์ LSTM ที่คำนวณจากสมการ (8)–(12) ผลลัพธ์สำหรับชั้น 2 จะถูกสร้างสมการ (13) ผลลัพธ์สุดท้ายจะเป็นการทำนายกิจกรรมของหน้าต่างขนาด  $T$



ภาพ 21 การทำงานของ Lightweight RNN–LSTM Preeti et al. (2020)

ชุดข้อมูลสาธารณะที่ใช้มาจากเซนเซอร์วัดความเร่งสามแกน โดยมี 6 กิจกรรม จาก 29 คน กิจกรรมได้แก่ การเดิน การเดินขึ้นบันได การเดินลงบันได การวิ่ง การยืน และการนั่ง ซึ่งแต่ละคนทำกิจกรรมที่แตกต่างกันโดยถือโทรศัพท์มือถือไว้ในกระเป๋าบริเวณขาด้านหน้า สุ่มตัวอย่างอัตราคงที่ 20 Hz แสดงดังตาราง

### ตาราง 3 อธิบายชุดข้อมูล WISDM

Total Number of Samples: 1,098,207		
Total Number of Subjects: 29		
Activity	Samples	Percentage
Walking	4,240,400	38.6%
Jogging	3,420,277	31.2%
Upstairs	1,220,869	11.2%
Downstairs	1,000,427	9.1%
Sitting	59,939	5.5%
Standing	48,397	4.4%

ในการทดสอบนี้ได้นำไปทดสอบบนราสเบอร์รี่พายสาม (Raspberry Pi3) โดยผลการทดสอบของแบบจำลองให้ความแม่นยำถึง 95.78% แบบจำลองมีขนาดเบา ลดข้อจำกัดด้านทรัพยากร เหมาะสำหรับอุปกรณ์ขนาดเล็ก แต่ในงานวิจัยนี้พัฒนาแบบจำลองโดยใช้ข้อมูลจากเซนเซอร์วัดความเร่งเพียงชนิดเดียว อาจมีประสิทธิภาพไม่ดี หากนำไปใช้กับข้อมูลชุดกิจกรรมอื่น ๆ ที่ใช้เซนเซอร์ชนิดอื่นได้ และไม่สามารถรู้จำกิจกรรมที่มีความซับซ้อนมากได้

#### 4. Human Activity Recognition Based on Wearable Sensor Using Hierarchical Deep LSTM Networks

จากงานวิจัย (Wang LuKun และLiu RuYue, 2020) ได้นำเสนอการรู้จำกิจกรรมของมนุษย์โดยใช้เซนเซอร์บนอุปกรณ์สวมใส่โดยใช้โครงข่ายความจำระยะสั้นแบบยาวตามลำดับชั้น (Hierarchical deep long short term : H-LSTM) ใช้ชุดข้อมูลสาธารณะ 2 ชุด ได้แก่ UCI และ HHAR ในแบบจำลองนี้ได้ใช้ความถี่ของเวลาในการแยกคุณลักษณะ กล่าวคือ ใช้หลักการสังเกตสัญญาณโดยดูที่โดเมนเวลา และโดเมนความถี่ เพื่อแยกคุณลักษณะ โดยเลือกจากส่วนเบี่ยงเบนมาตรฐาน ความถี่ ความไม่สมมาตร ค่าเฉลี่ยกำลังสอง และค่าเฉลี่ยความผิดพลาดสัมบูรณ์ มีขั้นตอนการทำงานอยู่ 4 ขั้นตอน ดังนี้ การเก็บข้อมูล และการเตรียมข้อมูลให้พร้อมใช้งาน, การเลือก และแยกคุณลักษณะเฉพาะ, การฝึกฝนแบบจำลอง และขั้นตอนสุดท้ายจำแนกประเภทกิจกรรมของมนุษย์ ฟังก์ชันในการอัปเดตเซลล์ของ LSTM แสดงดังสมการ

$$\bar{c}^{(t)} = \tan(W_{xc}x^{(t)} + W_{hc}h^{(t-1)} + b_c) \tag{14}$$

$$i^{(t)} = \sigma(W_{xi}x^{(t)} + W_{hi}h^{(t-1)} + W_{ci}c^{(t-1)} + b_i) \tag{15}$$

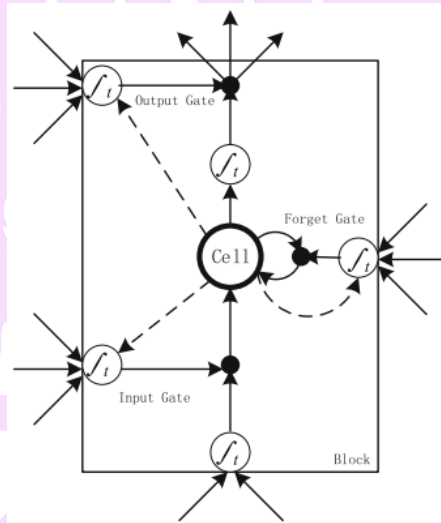
$$c^{(t)} = \delta(f^{(t)}, c^{(t-1)}) + \delta(i^{(t)}, c^{-^{(t)}}) \tag{16}$$

โดยที่  $x^t$  คือข้อมูลขาเข้า

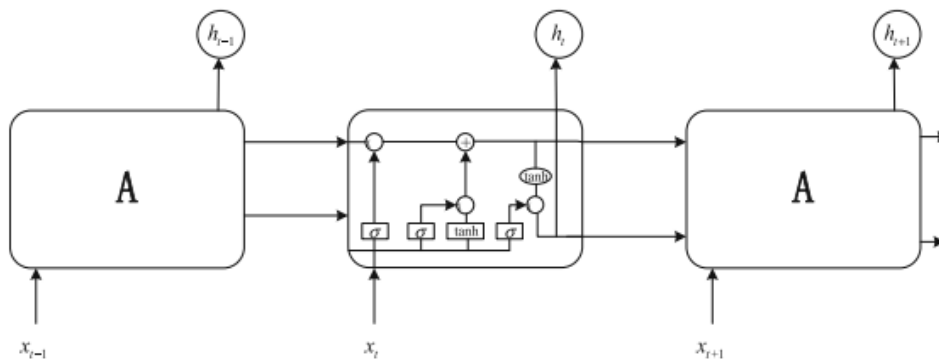
$h^t$  คือข้อมูลขาออก

$\delta$  คือฟังก์ชัน sigmoid

$\tan$  คือฟังก์ชัน tan



ภาพ 22 โครงสร้างในการอัปเดตเซลล์ LSTM



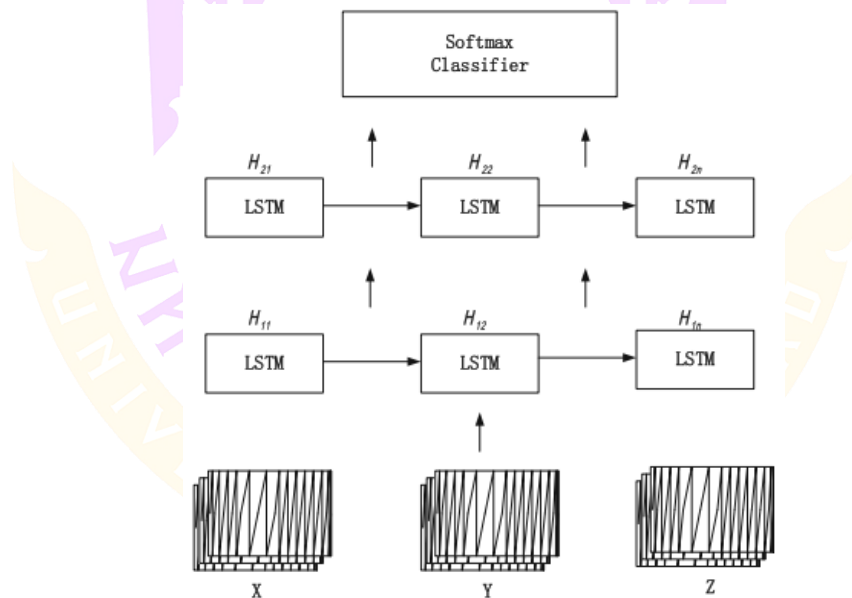
ภาพ 23 โครงสร้างของ LSTM

จากภาพโครงสร้างของ LSTM จะมีสมการทางคณิตศาสตร์ดังสมการ

$$o^t = \sigma(W_{xo}x^t + W_{ho}h^{t-1} + W_{co}c^{t-1} + b_o) \quad (17)$$

$$\bar{c}^{-(t)} = \delta(o^{(t)}, \tan(c^t)) \quad (18)$$

โดยที่  $W_{xc}, W_{hc}$  คือ weight ของข้อมูลขาเข้าและขาออก ณ เวลาก่อนหน้า  
 $\bar{c}^t$  คือค่าของตัวเลือกในหน่วยความจำ ณ เวลาปัจจุบัน  
 $b_c, b_i, b_o$  คือค่าพารามิเตอร์หน่วยความจำปัจจุบัน เกตขาเข้า และเกตขาออก  
 $i^{(t)}, f^{(t)}, o^{(t)}$  คือค่าของเกตขาเข้า เกตของ forget และเกตขาออก เกตขาเข้า  $i^{(t)}$   
 ประกอบด้วย 3 หน่วย หน่วยขาเข้า  $x^{(t)}$  หน่วยขาออก  $h^{(t-1)}$  และ  
 หน่วยความจำ  $c^{(t-1)}$   
 $\sigma$  คือ (0,1)  
 $\delta$  คือ ฟังก์ชันการคูณแบบจุดต่อจุด



ภาพ 24 โครงสร้างของแบบจำลอง H-LSTM

จากภาพโครงสร้างของแบบจำลอง H-LSTM ในชั้นข้อมูลขาเข้าจะนำข้อมูลจาก เซนเซอร์วัดความเร่ง ซึ่งเป็นเวกเตอร์สามมิติ จะถูกแยกออกเป็นหน้าต่างๆ หลาย ๆ หน้าต่าง

ด้วยความยาวของ  $N$  ข้อมูลขาเข้าเป็นอนุกรมเวลา 3 เมทริกซ์ จากนั้นจะนำไปเข้าชั้น hidden 2 ชั้น โดย  $H_{1n}$  คือ เซลล์ประสาท  $n$  ชั้นแรก และ  $H_{2n}$  เซลล์ประสาท  $n$  ชั้นที่ 2 และในชั้นสุดท้ายจะส่งข้อมูลขาออก

ในชุดข้อมูลที่ใช้ทดสอบ 2 ชุดข้อมูล ชุดแรกคือ UCI เป็นข้อมูลที่มาเซนเซอร์วัดความเร่ง และวัดการหมุน โดยสุ่มตัวอย่างอัตราคงที่ 50 Hz 1 กิจกรรมจะถูกแบ่งออกเป็นหน้าต่างละ 2.56 วินาที เป็น 128 หน้าต่าง ใช้วิธี time-frequency domain ตัดความถี่ที่ 0.3Hz แบ่งข้อมูลทดสอบเป็น 2 ส่วน ชุดฝึกฝน 75% และชุดทดสอบ 25% ชุดข้อมูล HHAR เป็นชุดข้อมูลที่แตกต่างกัน โดยประกอบด้วยชุดข้อมูล 2 ชุด จากสมาร์ทโฟน และสมาร์ทวอตช์ ทั้งหมด 6 กิจกรรม ดังนี้ การปั่นจักรยาน การนั่ง การยืน การเดิน การเดินขึ้นบันได และการเดินลงบันได แบ่งข้อมูลทดสอบเป็น 2 ส่วน ชุดฝึกฝน 70% และชุดทดสอบ 30%

งานวิจัยนี้นำวิธี time-frequency domain เข้ามาใช้เพื่อแยกคุณลักษณะของเวลา และความถี่ ซึ่งโดเมนเวลาจะใช้คุณสมบัติของสัญญาณความเร่งเป็นหลัก และโดเมนความถี่ใช้วิธีการแปลงฟูเรียร์ (Fourier transform: FFT) เพื่อแสดงถึงคุณลักษณะของความถี่ สมการของ time domain ที่ใช้แสดงดังสมการ

$$sd = \sqrt{\frac{1}{n} \sum_{i=1}^n (a_i - \bar{a})^2} \quad (19)$$

$$Kurtosis = \frac{\sum_{i=1}^n (a_i - \bar{a})^4}{(n-1) sd^4} \quad (20)$$

$$Skewness = \frac{\sum_{i=1}^n (a_i - \bar{a})^3}{(n-1) sd^3} \quad (21)$$

$$rms = \frac{\sqrt{\sum_{i=1}^n a_i^2}}{n} \quad (22)$$

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n |a_i - \bar{a}| \quad (23)$$

โดยที่  $a_i$  คือค่า  $i$  ในหน้าต่าง

$\bar{d}$  คือค่าเบี่ยงเบนสัมบูรณ์เฉลี่ยระหว่างหน้าต่างกับค่าเฉลี่ย

$\bar{a}$  คือค่าเฉลี่ยของข้อมูลกับหน้าต่าง

สมการทางคณิตศาสตร์ของ frequency domain แสดงดังสมการ

$$X_i(k) = \sum_{n=1}^n s_i n e^{-j2\pi kn/N}, k = 1, 2, \dots, N \quad (24)$$

$$C_i(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log X_i(e^{j\omega}) e^{j\omega n} d\omega \quad (25)$$

โดยที่  $X_i(n)$  คือเซปสตรัม (cepstrum)

$S_i, n$  คือข้อมูลในแถว  $n$  คอลัมน์  $i$

$s_j, n$  คือข้อมูลในแถว  $n$  คอลัมน์  $j$

จากผลการทดสอบแบบจำลอง H-LSTM มีความแม่นยำ 91.65% แต่ยังไม่ครอบคลุมกับกิจกรรมของมนุษย์ที่มีความซับซ้อนได้

### 5. LSTM-CNN Architecture for Human Activity Recognition

Xia et al. (2020) ได้เสนอสถาปัตยกรรมของโครงข่ายประสาทเทียมความจำระยะสั้นยาวและเครือข่ายประสาทเทียมสำหรับการรู้จำกิจกรรมของมนุษย์ โดยใช้ชุดข้อมูลสาธารณะ ได้แก่ UCI, WISDM และ OPPORTUNITY ซึ่งมีจำนวน 3 ชุดข้อมูล เพื่อใช้ในการประเมินผล แสดงดังตาราง

ตาราง 4 รายละเอียดของชุดข้อมูลที่ใช้ในงานวิจัย Xia et al. (2020)

Datasets	Activities	Sensors	S.Rate	Volunteers	Samples
UCI	6	A,G	50Hz	30	748,406
WISDM	6	A	20Hz	36	1,098,209
OPPORTUNITY	17	A,G,M,O,AM	30Hz	4	701,366

A=accelerometer, G=gyroscope, M=magnetometer, O=object sensor, AM=ambient sensor

ในขั้นตอนการเตรียมข้อมูลใช้ 3 วิธีดังนี้ 1. Linear interpolation เป็นอัลกอริทึมที่เพิ่มเติมส่วนข้อมูลบางส่วนที่ขาดหายไป เนื่องจากเป็นข้อมูลเซนเซอร์แบบไร้สาย ข้อมูลบางส่วนอาจสูญหายระหว่างกระบวนการเก็บรวบรวม 2. Scaling and Normalization เป็นวิธีการแปลงค่าขนาดใหญ่ให้อยู่ในช่วง 0-1 ดังสมการ

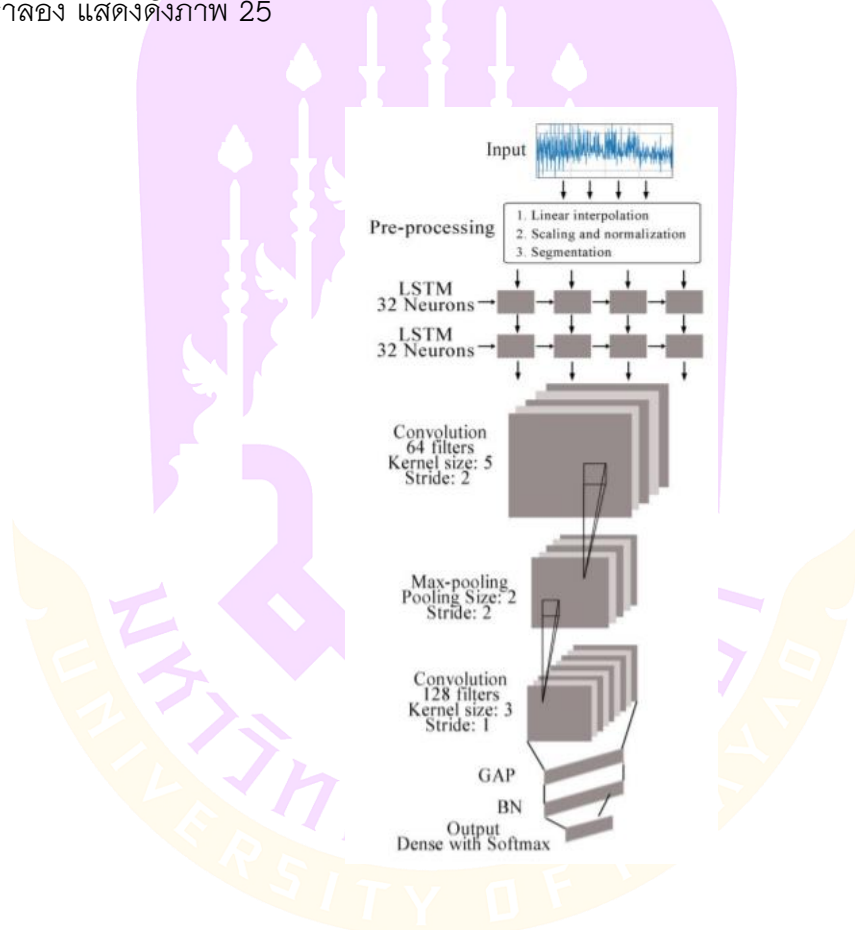
$$X_i = \frac{X_i - x_{i \min}}{X_{i \max} - x_{i \min}} \quad (i=1, 2, \dots, n) \quad (26)$$

โดยที่  $n$  คือจำนวนของสัญญาณ

$x_{i \max}, x_{i \min}$  คือค่าสูงสุดและต่ำสุด

สุดท้าย 3. Segmentation เพื่อรักษาความสัมพันธ์ระหว่างจุดข้อมูลในกิจกรรม จึงใช้การเลื่อนหน้าต่างที่อัตรา 50% ในชุดข้อมูล UCI และ WISDM มีหน้าต่างเท่ากับ 128 ในชุดข้อมูล OPPORTUNITY มีการบันทึกข้อมูลอยู่ในช่วงเวลาสั้น ๆ จึงต้องใช้หน้าต่างเลื่อนสั้น ๆ เพื่อให้ได้ตัวอย่างเพิ่มเติม

งานวิจัยนี้ได้กล่าวว่าโครงข่ายประสาทเทียมแบบตื้นที่ออกแบบมาสำหรับแก้ปัญหาในข้อมูลที่เป็นลำดับชั้น (Recurrent neural network : RNN) รวมกับแบบจำลองโครงข่ายประสาทเทียมความจำระยะสั้นยาว (Long short-term memory : LSTM) การทำงานของแบบจำลอง แสดงดังภาพ 25



ภาพ 25 การทำงานของแบบจำลอง LSTM-CNN Kun et al. (2020)

จากภาพการทำงานของแบบจำลอง LSTM-CNN ได้มีโครงสร้างของ LSTM และ CNN โดยการนำ LSTM มาใช้ในการแก้ไขปัญหาคือข้อมูลที่เป็นลำดับ ซึ่งข้อมูลจะถูกป้อนเข้าไป โดยผ่าน 2 ชั้นของ LSTM แต่ละชั้นจะมีเซลล์หน่วยความจำ 32 เซลล์ และจะถูกส่งไปยังเกตต่าง ๆ เพื่อควบคุมการทำงานของเซลล์ การเปิดใช้งานของแต่ละเซลล์ แสดงดังสมการ

$$h_t = \sigma(w_{i,h} \cdot x_t + w_{h,h} \cdot h_{t-1} + b) \quad (27)$$

โดยที่  $h_t, h_{t-1}$  คือการเปิดใช้งาน ณ เวลา  $t$  และ  $t-1$

$\sigma$  คือฟังก์ชันการเปิดใช้งานที่ไม่เป็นเชิงเส้น

$w_{i,h}$  คือ weight ของ hidden ขาเข้า

$w_{h,h}$  คือ weight ของ hidden

$b$  คือ hidden bias vector

ขาออกของ LSTM มีสามมิติ (samples, time step, input dimension) และในขั้นต่อไปเป็นขั้นของ Convolution และ Pooling โดยใช้ฟังก์ชันของ non-linear แสดงดังสมการ

$$a_{i,j} = f\left(\sum_{m=1}^M \sum_{n=1}^N w_{m,n} \cdot x_{i+mj+n} + b\right) \quad (28)$$

โดยที่  $a_{i,j}$  คือการเปิดใช้งาน

$w_{m,n}$  คือ weight ของ  $m \times n$  ของ kernel

$x_{i+mj+n}$  คือ activation ของเซลล์ประสาทส่วนบนที่เชื่อมต่อกับเซลล์

ประสาท (i,j)

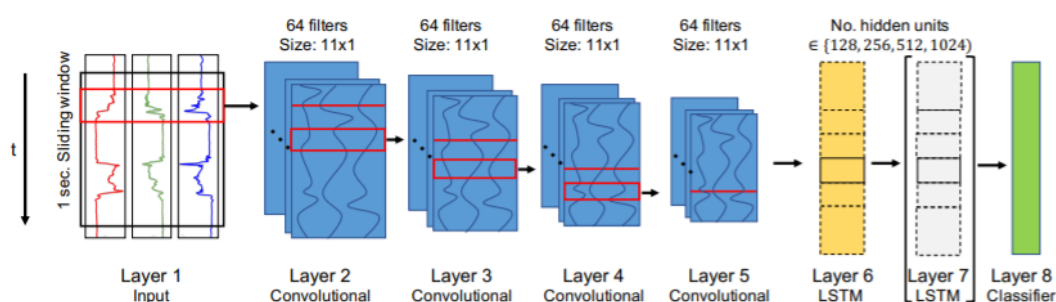
$b$  คือ ค่า bias

$f$  คือฟังก์ชัน non-linear

ในชั้น convolutional มี 64 convolutional ใช้สำหรับแยกคุณลักษณะ kernel  $1 \times 5$  เท่ากับ 128 เพื่อทำการสกัดคุณสมบัตินี้ ชั้นที่ 2 มี 128 convolutional kernel  $1 \times 3$  ในชั้น max-pooling จะอยู่ระหว่าง ชั้น convolutional ซึ่งจะทำหน้าที่ลดพารามิเตอร์ และกรองสัญญาณรบกวน ชั้นต่อไปหลังจากชั้น convolutional จะเป็นชั้น Global Average pooling หรือ GAP โดยจะทำการแปลง feature map เป็น 1 มิติ และใช้พารามิเตอร์น้อย ต่อไปส่งค่าไปยังชั้น Batch normalization (BN) เป็นชั้นที่ช่วยให้แบบจำลองฝึกฝนได้เร็วขึ้น ชั้นสุดท้ายชั้น softmax เป็นชั้นความน่าจะเป็นของแต่ละกิจกรรม แบบจำลองนี้ให้ค่าความแม่นยำที่สูงดังนี้ 95.80% 95.75% และ 92.63% จากชุดข้อมูล UCI, WISDM และ OPPORTUNITY ตามลำดับ งานวิจัยนี้สามารถแยกคุณสมบัตินี้ของกิจกรรมแบบอัตโนมัติ โดยใช้พารามิเตอร์ที่น้อยลง และมีความแม่นยำสูงขึ้น แต่เหมาะกับการประมวลผลในช่วงเวลาสั้น ๆ เท่านั้น

## 6. Improving Deep Learning for HAR with shallow LSTMs

Bock Marius และคณะ (2021) ได้เสนอการปรับปรุงการเรียนรู้เชิงลึกสำหรับการรู้จำกิจกรรมของมนุษย์ด้วยโครงข่ายประสาทแบบตื้น ใช้ชุดข้อมูลสาธารณะ 5 ชุด ได้แก่ Opportunity, Wetlab, RWHAR, SBHAR, และ HHAR จากการกล่าวโดย Chen et al. (2021) และจากการค้นพบของ Andrej Karpathy, Justin Johnson และ Li Fei-Fei (2015) ต้องมี LSTM อย่างน้อย 2 ชั้น เพื่อใช้จัดการกับข้อมูลลำดับ ได้มีการนำเสนอให้เปลี่ยนเป็นชั้นเดียว โดยแบบจำลองที่เสนอเป็นการรวมกันระหว่าง Convolutional และ Long Short-Term Memory (LSTM) ใช้แนวคิดของ Ordóñez Francisco Javier Roggen Daniel (2016) จากการรวมชั้นเครือข่ายทั้งสองแบบเข้าด้วยกัน เพื่อรวมชั้นเลเยอร์ที่มีการเกิดซ้ำซ้อน แสดงดังภาพ 26



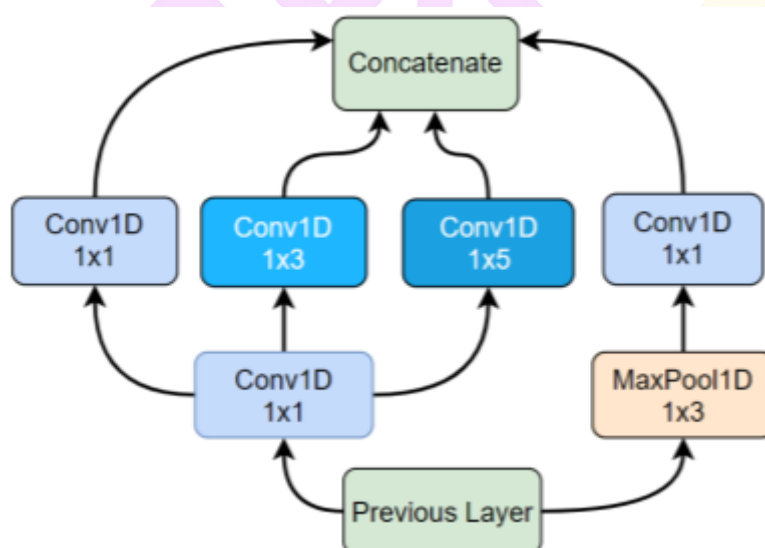
ภาพ 26 การรวมชั้นเครือข่าย Convolutional และ Long Short-Term Memory เข้าด้วยกัน (Ordóñez Francisco Javier และ Roggen Daniel, 2016)

จากภาพ 26 ใช้วิธีควบคุมการไหลของข้อมูลแบบ sliding window ซึ่งมีหน้าต่าง 1 วินาที ทับซ้อนกัน 60% เพื่อให้ได้ชุดไฮเปอร์พารามิเตอร์ที่เหมาะสม โดยใช้ พารามิเตอร์ Adam, weight decay ( $1e^{-6}$ ), learning rate ( $1e^{-4}$ ) ใช้ cross-entropy loss ในการคำนวณค่า error เพื่อเรียนรู้ชุดข้อมูลที่ไม่สมดุล ได้แก่ Wetlab และไม่ได้ทำ hyper tuning บนชุดข้อมูล RWHAR SBHAR และ HHAR โดยในชุดข้อมูล HHAR สุ่มตัวอย่างที่ 100 Hz จึงเพิ่มขนาดของตัวกรองขึ้นสองเท่าเป็น 21 เพื่อรักษาความสัมพันธ์ระหว่างตัวกรอง และ ขนาดของหน้าต่าง ทำให้อข้อมูลมีปริมาณเท่ากัน ชุดข้อมูล Opportunity ใช้ไฮเปอร์พารามิเตอร์ เพื่อเปรียบเทียบความแตกต่างของกิจกรรม เปลี่ยนขนาดของหน้าต่าง และใช้หน้าต่าง 0.5 วินาที ทับซ้อนกัน 50% โดยรวมผลรวมของสถาปัตยกรรมหนึ่งชั้นมีประสิทธิภาพเหนือกว่าสถาปัตยกรรมสองชั้น จาก

ชุดข้อมูล Opportunity, Wetlab, HHAR and SBHAR โดยชุดข้อมูล Opportunity ที่ใช้ชั้น hidden layer 1024 ได้ผลลัพธ์เทียบเท่ากับสถาปัตยกรรมดั้งเดิม การลดเลเยอร์ชั้นสองออกไปสามารถลดจำนวนพารามิเตอร์ลงได้ และลดเวลาในการฝึกฝนแบบจำลองโดยเฉลี่ย 38 % ขึ้นไป สำหรับเครือข่ายที่ใหญ่โดยเฉลี่ย 48 %

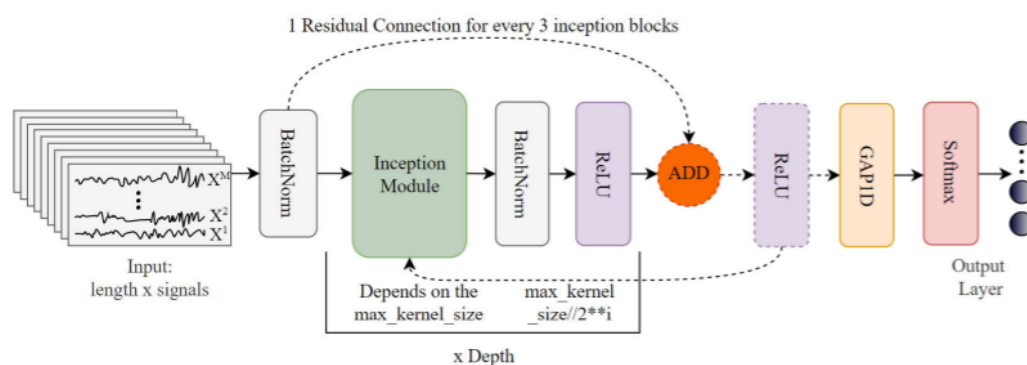
### 7. iSPLInception: An Inception-ResNet Deep Learning Architecture for Human Activity Recognition

จากงานวิจัย (Ronald Mutegeki, Poulouse Alwin และ Han Dong Seog, 2021) ได้นำเสนอสถาปัตยกรรมการเรียนรู้เชิงลึกสำหรับการรู้จำกิจกรรมของมนุษย์โดยใช้แบบจำลองของกุ๊กกึ่ง ให้ชื่อแบบจำลองว่า iSPLInception งานวิจัยนี้ได้ใช้ชุดข้อมูลสาธารณะทั้งหมด 4 ชุดในการประเมินผล ได้แก่ UCI, Opportunity, Daphnet freezing of gait และ PAMAP2 ผู้วิจัยได้อ้างถึง Szegedy Christian และคณะ (2017) ที่ได้มีการปรับปรุงงานวิจัยของเขาตลอดมาโดยได้แนะนำองค์ประกอบสำคัญที่ทำให้แบบจำลองขนาดใหญ่มากมีความสามารถในการนำมารวมกันที่เร็ว และทำงานได้ดี และการทำงานต่อจากนั้นมาจากการนำเสนอของ (He Kaiming และคณะ, 2016) ได้สร้างเครือข่ายที่กว้าง และลึก ทำให้มีประสิทธิภาพในการคำนวณที่น้อย สามารถนำมารวมกันเร็วมาก และยังสามารถปรับขนาดได้ โดยขั้นตอนที่ได้กล่าวมาข้างต้น แสดงดังภาพ 27



ภาพ 27 โมดูล Inception ของแบบจำลอง iSPLInception

จากภาพโมดูล Inception ในแบบจำลอง iSPLInception แต่ละโมดูลได้ทำการรันจาก convolution ที่มีขนาดของช่องสัญญาณรับเข้าของข้อมูล (Kernel) ที่ต่างกัน หลังจากนั้นได้เชื่อมต่อข้อมูลไปที่เลเยอร์ชั้นต่อไปทันที เริ่มจากข้อมูลในชั้นก่อนหน้านี้ เชื่อมต่อกับ Conv1D 1\*1 และ MaxPool1D 1\*3 ชั้นแรกของ Conv1D โดยในชั้นของ Conv1D 1\*1 จะถูกเชื่อมกับโหนดต่อไป ดังนี้ 1\*1, 1\*3 และ 1\*5 และ MaxPool1D เชื่อมต่อกับ 1\*1 Conv1D และชั้นต่อไปจาก Conv1D 1\*1 เชื่อมกับโหนด และ Conv1D 1\*3 1\*5 ทั้งสองจะถูกเชื่อมกับโหนด Conv1D ของ MaxPool1D ก็ถูกเชื่อมเช่นเดียวกัน โดยวิธีข้างต้นเป็นการลดขนาดของมิติในชั้นเลเยอร์ เพื่อให้ประหยัดค่าใช้จ่ายในการทำงาน



ภาพ 28 แบบจำลอง iSPLInception

จากภาพแบบจำลอง iSPLInception แสดงการทำงานของแบบจำลอง iSPLInception เริ่มจากข้อมูลขาเข้าเป็นกลุ่มของสัญญาณที่มีความยาวเท่ากัน (window size) ข้อมูลขาเข้ามีลักษณะของข้อมูลเป็นอาร์เรย์ โดยจำนวนตัวอย่างต่อลำดับของข้อมูลขาเข้า  $S$  เป็นดังนี้

$$S = \text{winsize} \times \text{samplingrate} \quad (29)$$

จากสมการ แสดงรูปแบบของข้อมูลขาเข้า โดย winsize คือขนาดของแต่ละหน้าต่างเป็นวินาที และ samplingrate คือ Hz เช่น จากชุดข้อมูล UCI มีสัญญาณจากเซนเซอร์ 6 ตัว สุ่มตัวอย่างที่ 50 Hz มีการสร้างหน้าต่างทับซ้อนกัน (overlapping window) 2.56 วินาที ซึ่งแปลงเป็น 128 ตัวอย่างข้อมูลต่อหน้าต่างต่าง เพื่อลดผลกระทบจากการใช้หน้าต่าง อาจเกิดเหตุการณ์ที่ข้อมูลบางส่วนหายไปเพราะขอบมีค่าเข้าใกล้ศูนย์ โดยสมการนี้ได้ช่วยแก้ปัญหาการจำแนกอนุกรมเวลาในชุดข้อมูล ชั้นตอนต่อไปหลังจากรับข้อมูลขาเข้า จะถูกส่งต่อไปยังชั้นที่ปรับค่าให้

มีขนาดเหมาะสม โดยเทียบจากค่าเฉลี่ย และค่าเบี่ยงเบนมาตรฐาน (BatchNorm) ซึ่งเป็น previous layer ในภาพ 28 ผลลัพธ์จะถูกส่งต่อไปยังชั้น BatchNorm อีกชั้นหนึ่งก่อนถึงจะถูกส่งไปที่ฟังก์ชันเส้นตรงที่ถูกปรับค่าให้เทรนแบบจำลองได้เร็วมากขึ้น (ReLU) จำนวนชั้นของโมดูล Inception, BatchNorm และ ReLU ถูกกำหนดโดยไฮเปอร์พารามิเตอร์มีความลึกในกรอบของ x Depth แบบจำลองนี้สามารถปรับขนาดได้ โดยมีบล็อกเริ่มต้นจำนวนมาก เพื่อให้แบบจำลองนี้กว้าง และลึก โดยไม่ทำให้ขนาดของแบบจำลองเพิ่มขึ้น จากทุก ๆ 3 บล็อก จะเพิ่มการเชื่อมต่อที่เหลือจากชั้นข้อมูลเข้าของชั้นแรก ผลลัพธ์ของผลรวมในการประมวลผลทั้งหมดจะเพิ่มเข้าไปในเลเยอร์ โดยใช้ ReLU การเชื่อมต่อที่เหลือจะขึ้นอยู่กับความลึกของไฮเปอร์พารามิเตอร์ที่เสนอ การใช้ RuLU ครั้งที่สองจะขึ้นอยู่กับว่าบล็อกได้มีการเชื่อมต่อกับที่ชั้นเหลือหรือไม่ ชั้นต่อไปใช้ 1-dimensional global average pooling (GAP1D) จากงานที่นำเสนอโดย Ismail Fawaz et al. (2020) เป็นการสุ่มตัวอย่างจากข้อมูลขาเข้าด้วยการหาค่าเฉลี่ยของข้อมูลขาเข้า ซึ่งจะได้ออกมาเป็นค่าเฉลี่ย 1 ค่า จากนั้นจะส่งค่าที่ได้ไปที่ชั้นต่อไปโดยชุดข้อมูล UCI PAMAP2 และ Opportunity ใช้ Softmax เพื่อคำนวณความน่าจะเป็นที่อยู่ระหว่าง 0-1 ส่วนชุดข้อมูล Daphnet ใช้ Sigmoid เนื่องจาก Softmax ไม่สามารถคำนวณข้อมูลขาเข้าที่มีค่าเป็นบวก ลบ ศูนย์ หรือมากกว่า 1 ได้ และสุดท้ายจะส่งข้อมูลขาออกของหมวดหมู่กิจกรรมของมนุษย์ออกมา การประเมินผลจากชุดข้อมูล 4 ชุด โดยเปรียบเทียบกับแบบจำลอง 6 แบบ ดังนี้ a vanilla long short-term memory (vLSTM), stacked long short-term memory (sLSTM), convolutional neural network (CNN), CNN-LSTM และ bidirectional long short-term memory (BiLSTM) networks ผลลัพธ์ที่ได้ แสดงดังตาราง 5

ตาราง 5 ผลลัพธ์จากจำลอง iSPLInception กับแบบจำลองอื่น ๆ จากชุดข้อมูล 4 ชุด

Models	Datasets			
	UCI	Opportunity	DAPHNET	PAMAP2
CNN	91.67	80.24	92.97	85.78
CNN-LSTM	94.48	81.41	92.97	88.37
vLSTM	90.80	76.79	93.22	85.52
sLSTM	91.82	80.82	87.65	86.97
BiLSTM	93.91	79.90	92.41	86.97
iSPLInception	95.09	88.14	93.52	89.09

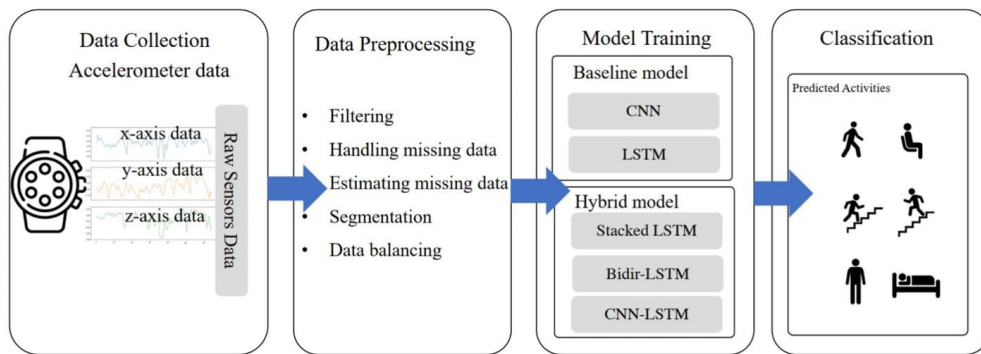
จากตาราง 5 จะเห็นได้ว่า iSPLInception มีความแม่นยำสูงกว่าแบบจำลองอื่น ๆ ที่นำมาเปรียบเทียบ จากทุกชุดข้อมูล จากชุดข้อมูล UCI, DAPHNET, PAMAP2 และ Opportunity โดยเรียงจากชุดข้อมูลที่มีความแม่นยำจากมากไปน้อย ตามลำดับ ผลลัพธ์ที่ได้ใช้ทรัพยากรในการคำนวณน้อยลง และมีประสิทธิภาพสูงกว่าสถาปัตยกรรมครั้งล่าสุดที่มีการนำเสนอสำหรับแก้ปัญหาการรู้จำของมนุษย์ แต่ยังมีข้อจำกัดในเรื่องของข้อมูล เนื่องจากความไม่สอดคล้องกันของชุดข้อมูลทำให้บางกิจกรรม เช่น ชุดข้อมูล UCI การนั่ง และการยืน ชุดข้อมูล Opportunity การเปิดลิ้นชัก และการปิดลิ้นชัก ชุดข้อมูล DAPHNET การเดินติดขัด และชุดข้อมูล PAMAP2 การยืน การดูคู่มือ และการรีดผ้า กิจกรรมของแต่ละชุดข้อมูลที่กล่าวมานั้นมีลักษณะที่ใกล้เคียงกันมาก จึงทำให้ถูกจำแนกได้ไม่ดี และส่งผลกระทบต่อข้อมูลที่ยังไม่ผ่านการฝึกฝนในแบบจำลอง โดยแบบจำลองทำงานได้ไม่ดีกับข้อมูลที่ไม่ได้รับการฝึกฝน



### บทที่ 3

#### วิธีดำเนินการวิจัย

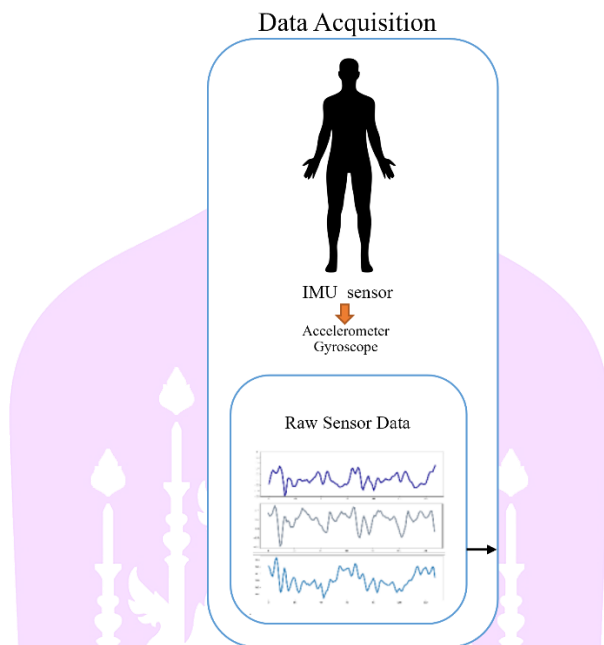
การศึกษาครั้งนี้ได้นำเสนออัลกอริทึมการเรียนรู้จำกิจกรรมในชีวิตประจำวันของมนุษย์ การทำงานของระบบการเรียนรู้จำกิจกรรมในชีวิตประจำวันของมนุษย์ ประกอบไปด้วย ขั้นตอนการประมวลผลเบื้องต้น ขั้นตอนการเตรียมข้อมูล และขั้นตอนการเรียนรู้จำ การทำงานของระบบเรียนรู้จำกิจกรรมในชีวิตประจำวันของมนุษย์ แสดงได้ดังภาพ



ภาพ 29 แสดงขั้นตอนกระบวนการเรียนรู้จำกิจกรรมในชีวิตประจำวันของมนุษย์

จากภาพ 29 รายละเอียดการทำงานในแต่ละขั้นตอนของระบบเรียนรู้จำกิจกรรมในชีวิตประจำวันของมนุษย์ ขั้นตอนแรกเป็นการเก็บรวบรวมข้อมูลขั้นตอนที่สองเป็นขั้นตอนการประมวลผลเบื้องต้น ขั้นตอนที่สามการเตรียมข้อมูล ขั้นตอนสุดท้ายเป็นขั้นตอนการฝึกฝนแบบจำลอง

## การเก็บข้อมูล (collection)



ภาพ 30 แสดงภาพรวมการเก็บรวบรวมข้อมูล

จากภาพ 30 ข้อมูลที่นำมาใช้ในการรู้จำกิจกรรมในชีวิตประจำวันของมนุษย์ในงานวิจัยนี้ จะใช้ชุดข้อมูลสาธารณะที่ประกอบด้วยข้อมูลทางกายภาพ ซึ่งเป็นข้อมูลมาจากแหล่งทุติยภูมิที่มีผู้อื่นเก็บรวบรวม และบันทึกไว้ เพื่อนำมาใช้ในการวิเคราะห์ในงานต่าง ๆ ได้อย่างหลากหลาย โดยชุดข้อมูลที่นำมาใช้ทดสอบในงานวิจัยนี้มีได้ทดลองกับชุดข้อมูลสาธารณะ ดังนี้

1) PAMAP2 เป็นชุดข้อมูลสาธารณะจาก UCI ประกอบด้วยกิจกรรมทางกายภาพ 18 แบบ เก็บข้อมูลจากอาสาสมัคร 9 คน สวมเซนเซอร์ IMU และเครื่องวัดอัตราการเต้นของหัวใจในการสวมตัวอย่าง 100 Hz โดยตำแหน่งของเซนเซอร์มี 3 ที่ เหนือข้อมือข้างที่ถนัด หน้าอก และข้อเท้าข้างที่ถนัด เครื่องมือใช้วัดอัตราการเต้นของหัวใจมีความถี่ในการสวมตัวอย่างประมาณ 9 Hz มีกิจกรรมดังนี้ 1 นอน นั่ง ยืน เดิน วิ่ง ปั่นจักรยานเดินนอร์ดิก ดูทีวี คอมพิวเตอร์ทำงาน ขับรถ ขึ้นบันได บันได เครื่องดูดฝุ่น รีดผ้า ซักผ้า พับ ทำความสะอาดบ้าน เล่นฟุตบอล กระโดดเชือก และกิจกรรมอื่น ๆ (กิจกรรมชั่วคราว) แสดงดังตาราง 6 (Reiss Attila และStricker Didier, 2012)

## ตาราง 6 กิจกรรมในชุดข้อมูล PAMAP2

Activities of PAMAP2	
Lying down	Ironing clothes
Sitting	Vacuum cleaning
Standing	Nordic walking
Walking	Ascending stairs
Running	Descending stairs
Cycling	Jumping rope

2) Smart Physiotherapy Activity Recognition System 9-axis (SPARS9x Dataset) เป็นข้อมูลที่เก็บรวบรวมมาจากอาสาสมัครที่เข้าร่วมในโครงการจำนวน 20 คน ประกอบด้วย เพศชาย 8 คน เพศหญิง 12 คน อายุเฉลี่ย 25 ปี เก็บรวบรวมข้อมูลจากเซ็นเซอร์ 3 ตัว ได้แก่ เซ็นเซอร์วัดความเร่ง เซ็นเซอร์วัดการหมุน และเซ็นเซอร์วัดค่าความเข้มข้นแม่เหล็ก จากหัวเว่ยสมาร์ทวอตช์ อาสาสมัครแต่ละคนจะทำกิจกรรมที่กำหนดให้ซ้ำ ๆ จำนวน 20 ครั้ง และกิจกรรมต่าง ๆ ในชีวิตประจำวัน เป็นเวลาตามที่กำหนด แต่ละคนจะสวมอุปกรณ์พกพา หรือ อุปกรณ์ขนาดเล็กที่มีเซ็นเซอร์ฝังอยู่ เซ็นเซอร์ที่ฝังในอุปกรณ์ดังกล่าวเช่น เซ็นเซอร์วัดความเร่ง เซ็นเซอร์วัดการหมุน และเซ็นเซอร์วัดค่าแม่เหล็ก ในชุดข้อมูลจะมีป้ายกำกับ หรือ คำตอบของกิจกรรมระบุไว้ในแต่ละแถว เพื่อให้สามารถนำไปฝึกฝนแบบจำลองในลักษณะของการเรียนรู้แบบมีผู้สอน (Philip Boyer, Burns. และ Cari Whyne, 2020)

3) The w-HAR dataset เป็นชุดข้อมูลที่เก็บรวบรวมข้อมูลจากผู้เข้าร่วมจำนวน 22 คน (ผู้หญิง 8 คน และผู้ชาย 14 คน) โดยผู้เข้าร่วมจะดำเนินการทำกิจกรรมการเคลื่อนไหวต่าง ๆ ขณะสวมใส่ Invensense-9250 และ stretch sensors ซึ่งเป็นเซ็นเซอร์ที่ใช้ในการวัดแรงยึดเพื่อวัดท่าทางของร่างกายหรือการเคลื่อนไหว กิจกรรมทั้งหมด 8 กิจกรรม ได้แก่ กระโดด นอนลง นั่ง เดินลงบันได เดินขึ้นบันได ยืน เดิน และอยู่นิ่ง ดังแสดงดังตาราง 6 (Ganapati และคณะ, 2020)

ตาราง 7 กิจกรรมการเคลื่อนไหวในชีวิตประจำวันของชุดข้อมูลสาธารณะ w-HAR

Activity	Type
Lie Down	Static
Sit	Static
Stand	Static
Jump	Dynamic
Stairs Down	Dynamic
Stairs Up	Dynamic
Walk	Dynamic
Transition	Transition

ในชุดข้อมูลสาธารณะนี้เก็บรวบรวมข้อมูลจาก IMU บนข้อเท้าด้านขวาของผู้เข้าร่วม โดยใช้เซนเซอร์ 2 ตัว ดังนี้ เซนเซอร์วัดความเร่ง และเซนเซอร์วัดการหมุน โดยเก็บข้อมูลทั้ง 3 แกน (x,y,z) และข้อมูลจาก stretch sensor บนข้อเท้าขวาของผู้เข้าร่วม ข้อมูลจาก IMU ใช้ อัตรา 250 Hz และ stretch sensor ถูกบันทึกที่ 25 Hz ในแต่ละกิจกรรมผู้เข้าร่วมจะทำกิจกรรมตามเวลาที่กำหนดดังตาราง 7 และเวลาในการทำกิจกรรมต่าง ๆ แสดงดังตาราง 8

ตาราง 8 เวลาในการทำกิจกรรมการเคลื่อนไหวในชีวิตประจำวันของชุดข้อมูล  
สาธารณะ w-HAR

Experiment	Activity	Time
1	Stand	30 s
	Jump	3 times
	Stand	30 s
2	Stand	10 s
	Sit	30 s
	Stand	10 s
	Jump	3 times
3	Sit	30 s
	Stand	10 s
	Walk	40 steps
	Stand	Stand 10 s
4	Stand	10 s
	Jump	3 times
	Walk	40 steps
	Sit	20 s
5	Stand	10 s
	Sit	10 s
	Lie down	30 s
6	Sit	10 s
	Stand	10 s
	Walk down stairs	
7	Stand	10 s
	Stand	10 s
	Walk up stairs	
	Stand	10 s

4) MEx (Multi-modal Exercises) เป็นชุดข้อมูลโปรแกรมการออกกำลังกาย อาสาสมัครที่เข้าร่วมในโครงการจำนวน 30 คน อายุระหว่าง 18–24 ปี เก็บรวบรวมข้อมูลจาก เซ็นเซอร์ 4 ตัว เซ็นเซอร์วัดความดัน depth camera และเซ็นเซอร์วัดความเร่งของตัว 2 ตัว อาสาสมัครแต่ละคนจะทำกิจกรรมที่กำหนดให้ทั้งหมด 7 กิจกรรม ได้แก่ Knee-rolling, Bridging, Pelvic tilt, The Clam Repeated Extension in Lying, Prone punches, Superman ชุดข้อมูลประกอบด้วย 3 รูปแบบข้อมูลดังนี้ ข้อมูล Time series ข้อมูลวิดีโอ และข้อมูลเซ็นเซอร์ความดัน (Wijekoon, Wiratunga และ Cooper, 2019)

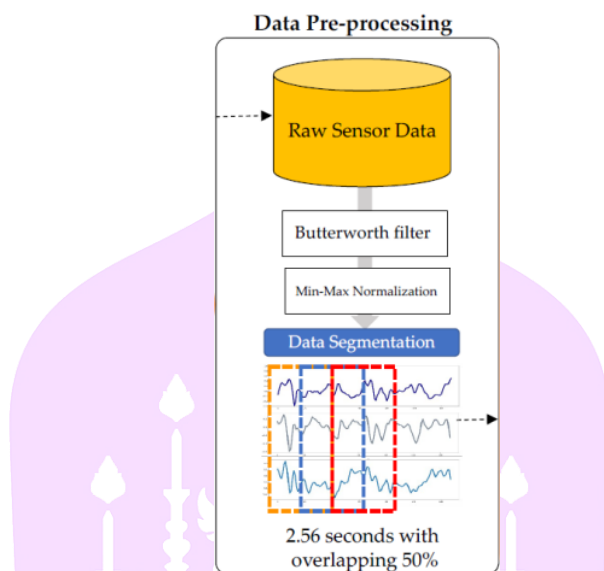
การอ่านข้อมูลจากไฟล์ .csv ในชุดข้อมูลสาธารณะต่าง ๆ แสดงดังภาพ 31

```
[ ] df = pd.read_csv('/content/drive/MyDrive/Dataset/motion_data_22_users.csv')
```

ภาพ 31 แสดงการใช้งาน pandas อ่านไฟล์นามสกุล csv



## การประมวลผลเบื้องต้น (processing)



ภาพ 32 แสดงภาพรวมการประมวลผลข้อมูล

จากภาพ 32 ขั้นตอนการประมวลผลเบื้องต้นจะประกอบไปด้วย การลดเสียงรบกวน การปรับปรุงจากนั้นจะจัดการโครงสร้างข้อมูลให้อยู่ในรูปแบบเดียวกัน ต่อมาจะทำการแบ่งกลุ่ม หรือจัดกลุ่มข้อมูล รายละเอียดการทำงานในขั้นตอนการประมวลผลเบื้องต้นผู้วิจัยได้ทดสอบหลายวิธีการโดยแบ่งได้ดังนี้

### 1) Noise Removing

**Median Filter** ตัวกรองที่นำมาลดสัญญาณรบกวน โดยจะคำนวณหาค่ากลางของค่าความเข้มสีของบริเวณโดยรอบจุดที่ต้องการ เป็นการนำเอาความเข้มแสงของจุดที่ตรงกันมาเรียงลำดับจากน้อยไปหามาก จากนั้นเลือกค่าที่อยู่ตรงกลางไปใช้ วิธีนี้ใช้เวลาในการคำนวณสูง แต่จะทำให้ไม่สูญเสียความคมชัด

**Butterworth แบบ Low-pass filter ลำดับที่ 3** การกรองสัญญาณ โดยยอมให้สัญญาณความถี่ตั้งแต่ 0 Hz จนถึงความถี่ที่กำหนดผ่านไปได้ โดยความถี่ที่สูงกว่าจะถูกลดทอนไปตามลำดับ ดังสมการ

$$\omega_c = \frac{1}{RC} = 2\pi fc \quad (30)$$

โดยที่  $\omega_c$  คือความถี่ตัดเชิงมุม (rad/s)

$fc$  คือความถี่ตัดเชิงเส้น (Hz)

$R$  คือความต้านทานไฟฟ้า ( $\Omega$ )

$c$  คือความจุของตัวเก็บประจุ (F)

ตัวอย่างการใช้งาน Butterworth แบบ Low-pass filter ลำดับที่ 3 จากไลบรารี scipy แสดงดังภาพ

```
[ ] def butterworthfilter(data):
    train_sig = data.copy()
    cutoff=50
    measurements=data.shape[0]
    time=0.02
    sampling_rate = measurements/time
    nyquist = sampling_rate*0.5
    wn = cutoff/nyquist
    b, a = butter(3, wn, btype='lowpass')
    sensor_data = ['Acc_x', 'Acc_z', 'Acc_y', 'Gy_x', 'Gy_y', 'Gy_z']
    for s in sensor_data:
        filtered_sig = ''
        filtered_sig = lfilter(b, a, data[s].values)
        train_sig[s] = filtered_sig
    return train_sig

[ ] from scipy.signal import butter, lfilter
    butterworth = butterworthfilter(df)
```

ภาพ 33 แสดงตัวอย่างการใช้งาน Butterworth แบบ Low-pass filter

## 2) Normalization

ข้อมูลที่ได้มามีความหลากหลาย ทั้งรูปแบบ ชนิด และ Scale ช่วงของข้อมูล ซึ่งในแบบจำลองหลาย ๆ ตัวไม่สามารถรับข้อมูลหลาย Scale ได้โดยตรง จึงต้องทำ Normalization ก่อนที่จะนำข้อมูลเข้าไปในแบบจำลอง

**Min-Max Normalization** การปรับปรุงโครงสร้างข้อมูลให้อยู่ในรูปแบบเดียวกัน เพื่อลดความซ้ำซ้อนของข้อมูล ซึ่งจะทำให้ขนาดข้อมูลลดลง และรักษาความถูกต้องของข้อมูล โดยการปรับช่วงข้อมูลให้อยู่ในช่วง [0,1] ด้วยการนำ feature ลบด้วยค่าที่น้อยที่สุด แล้วหารด้วยช่วงของข้อมูลนั้น ดังสมการ

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (31)$$

ตัวอย่างการใช้งาน Min-Max Normalization จากไลบรารี sklearn แสดงดังภาพ

```
[ ] from sklearn.preprocessing import MinMaxScaler
    scaler = MinMaxScaler()
    df.iloc[:,[1,-2]] = scaler.fit_transform(df.iloc[:,[1,-2]] )
```

ภาพ 34 แสดงตัวอย่างการใช้งาน Min-Max Normalization

**Standardization (Z-Score Normalization)** การปรับปรุงโครงสร้างข้อมูลให้อยู่ในรูปแบบเดียวกัน โดยการนำข้อมูลมาปรับให้ค่าเฉลี่ยมีค่าเท่ากับศูนย์ และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับหนึ่ง

$$x' = \frac{x - \bar{x}}{\sigma} \quad (32)$$

ตัวอย่างการใช้งาน Standardization จากไลบรารี sklearn แสดงดังภาพ

```
[ ] from sklearn.preprocessing import StandardScaler
    scaler = StandardScaler()
    df.iloc[:,[1,-2]] = scaler.fit_transform(df.iloc[:,[1,-2]] )
```

ภาพ 35 แสดงตัวอย่างการใช้งาน Standardization

### 3) Segmentation

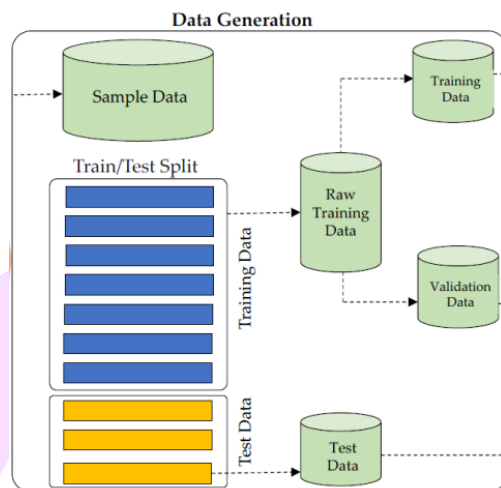
การแบ่งส่วนของข้อมูล แบ่งส่วนของข้อมูลด้วยค่าความเหมือนที่ทับซ้อนกัน 50% (Garcia-Gonzalez และคณะ, 2020) แสดงดังภาพ

```
[ ] def segment(data, window_size): # data is numpy array
    n = len(data)
    X = []
    y = []
    start = 0
    end = 0
    while start + window_size - 1 < n:
        end = start + window_size - 1
        if data[start][0] == data[end][0] and data[start][-1] == data[end][-1]:
            X.append(data[start:(end+1), 1:-2])
            y.append(data[start][0])
            start += window_size // 2 # 50% overlap
        else:
            while start + window_size - 1 < n:
                if data[start][0] != data[start+1][0]:
                    break
                start += 1
            start += 1
    return {'inputs': np.asarray(X), 'labels': np.asarray(y, dtype=int)}
```

ภาพ 36 แสดงตัวอย่างการแบ่งส่วนของข้อมูลที่ทับซ้อนกัน 50%



## การเตรียมข้อมูล (data generation)



ภาพ 37 แสดงภาพรวมการเตรียมข้อมูล

ในงานวิจัยนี้ได้สร้างข้อมูล เพื่อเตรียมข้อมูลให้พร้อมสำหรับนำไปใช้ฝึกฝน และทดสอบ การสร้างข้อมูลเป็นการนำข้อมูลที่ได้จากขั้นตอนการประมวลผลข้อมูลมาแบ่งเพื่อนำเข้าแบบจำลอง โดยใช้ K-Fold Cross Validation ซึ่งเป็นเครื่องมือในการแบ่งข้อมูลเป็นจำนวน K ส่วน ในการแบ่งข้อมูลแต่ละส่วนจะมาจากการสุ่มเพื่อให้ข้อมูลกระจายเท่า ๆ กัน การใช้เครื่องมือนี้สามารถใช้เปรียบเทียบข้อมูลชุดไหนดีที่สุดที่สุดในแบบจำลอง และเปรียบเทียบว่าแบบจำลองไหนดีกว่ากัน ในการศึกษาและวิจัยนี้ข้อมูลจะถูกแบ่งเป็น 3 ส่วน ส่วนแรกเป็นชุดที่จะนำข้อมูลเข้าไปฝึกฝนในแบบจำลอง (Training set) ส่วนที่สองเป็นชุดที่จะนำไปทดสอบหาค่าเพื่อวัดผลหลังจากปรับจูนแบบจำลองในแต่ละครั้งว่าแบบจำลองทำงานได้ดีเท่าใด (Testing set) ส่วนสุดท้ายเป็นส่วนที่ใช้ทดสอบแบบจำลองหลังจากเทรนแบบจำลองแล้ว (Validation set) ดังภาพ 38



ภาพ 38 การแบ่งชุดข้อมูลสำหรับการฝึกฝน การทดสอบ

ตัวอย่างการใช้งาน K-Fold Cross Validation แสดงดังภาพ

```

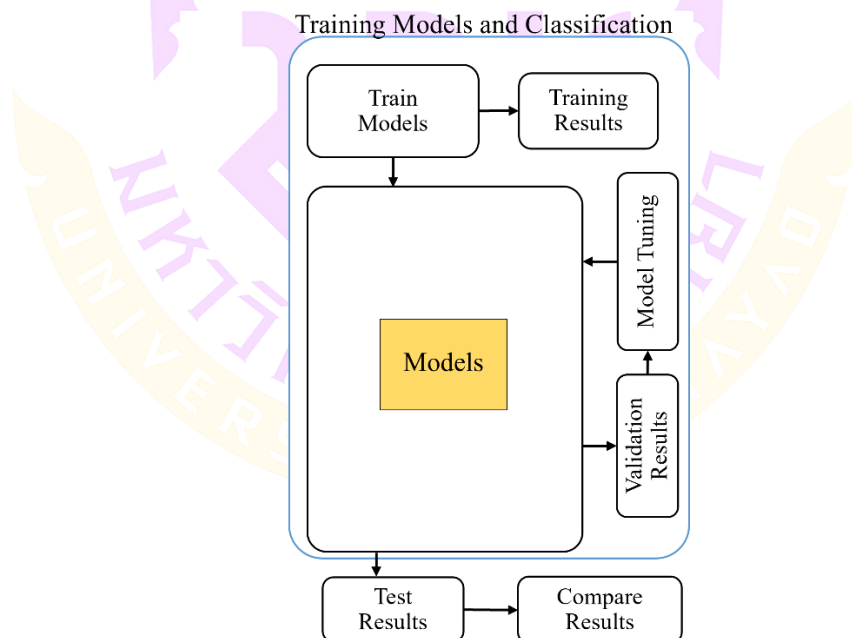
from sklearn.model_selection import KFold
kfold = KFold(n_splits=5, shuffle=True)

for train, test in kfold.split(X, y):
    K = len(set(y[train]))
    i = Input(shape=X[train][0].shape)
    x = Conv1D(32, 4, activation = 'relu')(i)
    x = MaxPooling1D()(x)
    x = Conv1D(64, 4, activation = 'relu')(i)
    x = MaxPooling1D()(x)
    x = Conv1D(128, 4, activation = 'relu')(i)
    x = MaxPooling1D()(x)
    x = Dropout(0.5)(x)
    x = Flatten()(x)
    x = Dense(K, activation = 'softmax')(x)
    model = Model(i,x)
    model.compile(optimizer = Adam(learning_rate=0.001),
                  loss = 'sparse_categorical_crossentropy',
                  metrics = ['accuracy'])
    history = model.fit(X[train], y[train], validation_data = (X[test], y[test]), epochs = 100, batch_size = 32, verbose=0 )

```

ภาพ 39 ตัวอย่างการใช้งาน K-Fold Cross Validation

การฝึกฝนแบบจำลอง และการจัดหมวดหมู่กิจกรรมของมนุษย์ (training model)

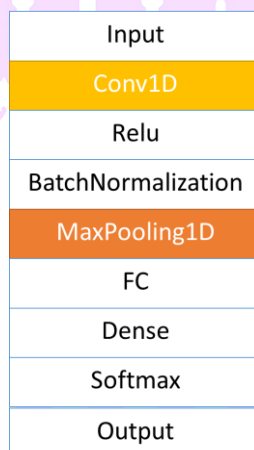


ภาพ 40 การฝึกฝนแบบจำลอง และจัดหมวดหมู่กิจกรรมของมนุษย์

ในขั้นตอนนี้จะเป็นขั้นตอนการนำข้อมูลที่ทำการเตรียมไว้มาทำการฝึกฝน และทดสอบ โดยแบบจำลองที่นำมาใช้ในการศึกษาและวิจัยครั้งนี้ ผู้วิจัยได้ทดสอบกับแบบจำลองโครงข่ายประสาทเทียมแบบคอนโวลูชัน แบบจำลองโครงข่ายประสาทเทียมแบบวนกลับ นอกจากนี้ยังมีการใช้เทคนิค Ensemble ในการสร้างแบบจำลอง การศึกษาและวิจัยครั้งนี้ได้ทดลองและออกแบบแบบจำลองจำนวนหลายตัว ดังนี้

### 1) CNN model

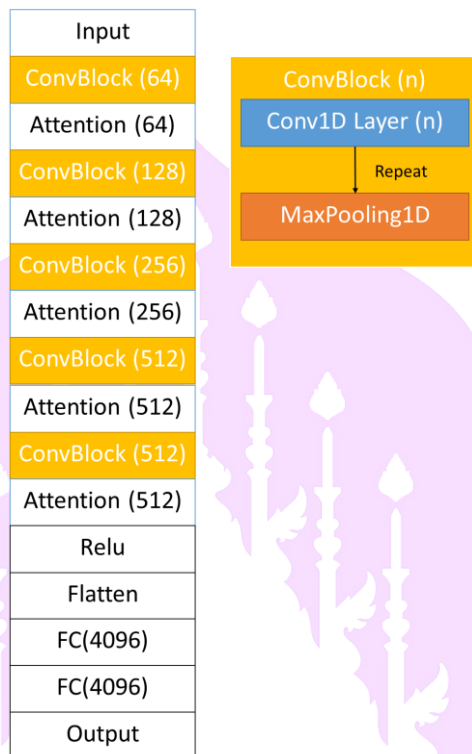
โครงสร้างของโครงข่ายประสาทเทียมแบบคอนโวลูชันในงานวิจัยนี้ได้ออกแบบแบบจำลองหลักแสดงดังภาพ



ภาพ 41 โครงสร้างหลักของโครงข่ายประสาทเทียมแบบคอนโวลูชันที่ใช้ในงานวิจัยนี้

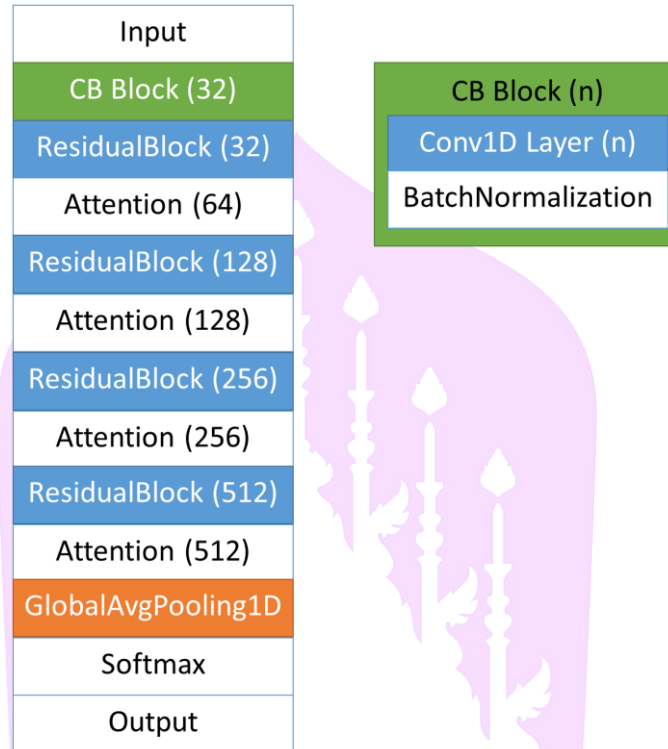
นอกจากนี้ทางผู้วิจัยทดสอบกับแบบจำลองน้ำหนักเบาอื่น ๆ ที่พัฒนาไว้แล้วมาทดสอบเพิ่มเติม ดังนี้

## 2) VGG16



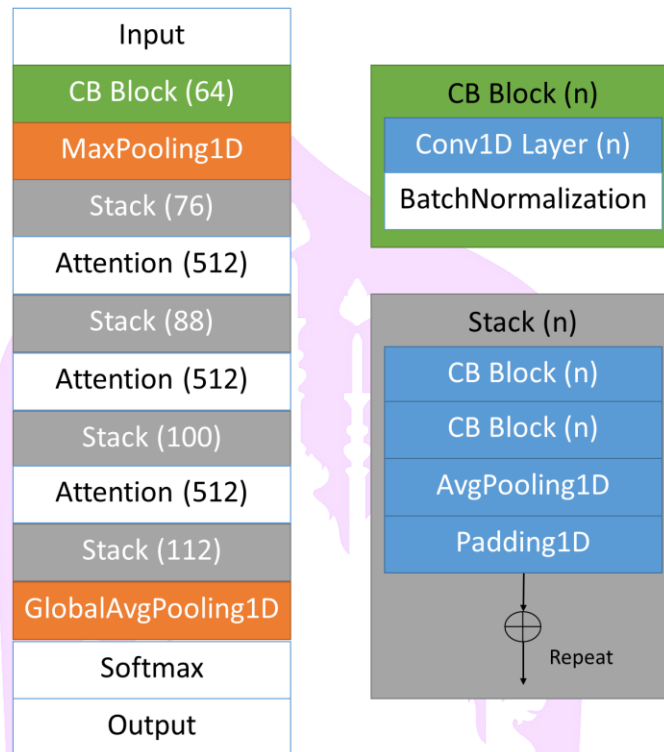
ภาพ 42 โครงสร้างแบบจำลอง VGG16

## 3) RESNET18



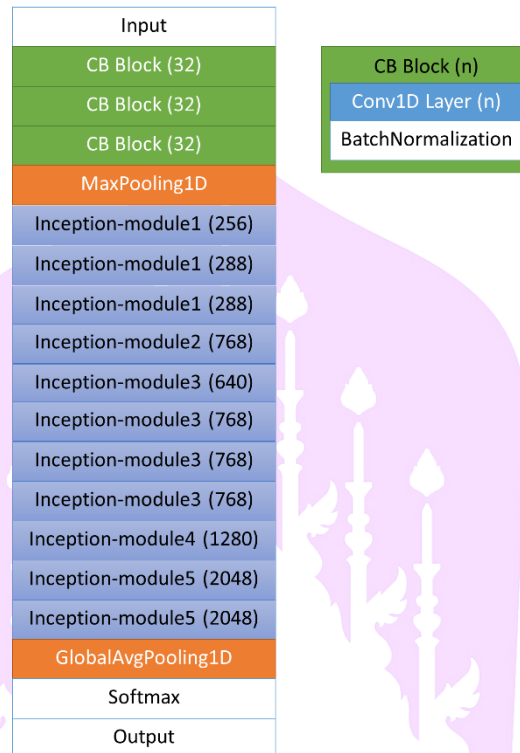
ภาพ 43 โครงสร้างแบบจำลอง RESNET18

## 4) PyramidNet18



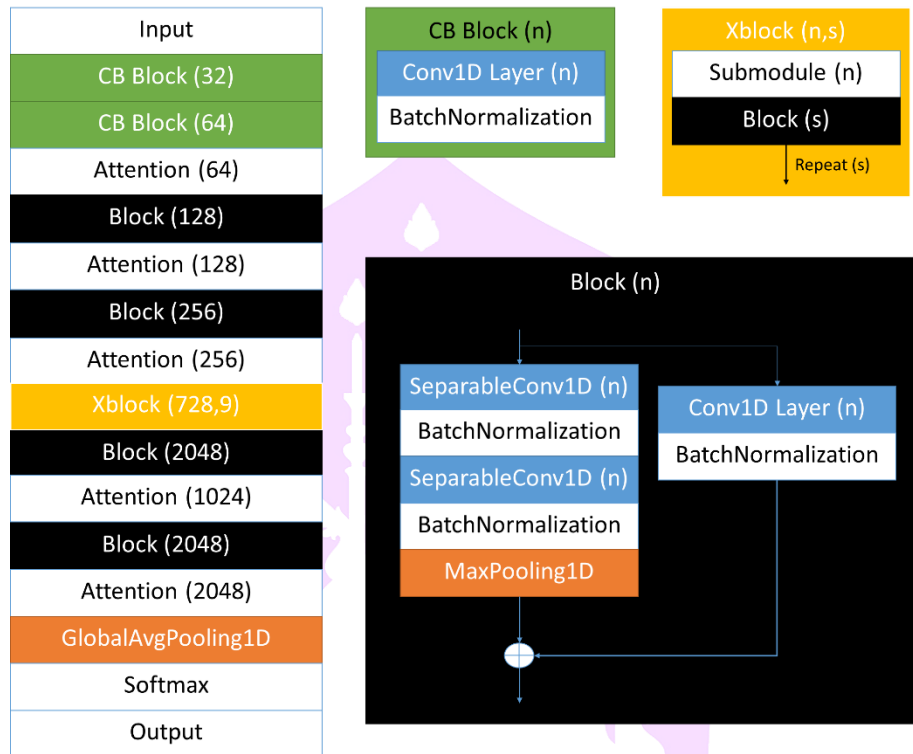
ภาพ 44 โครงสร้างแบบจำลอง PyramidNet18

## 5) Inception-V3



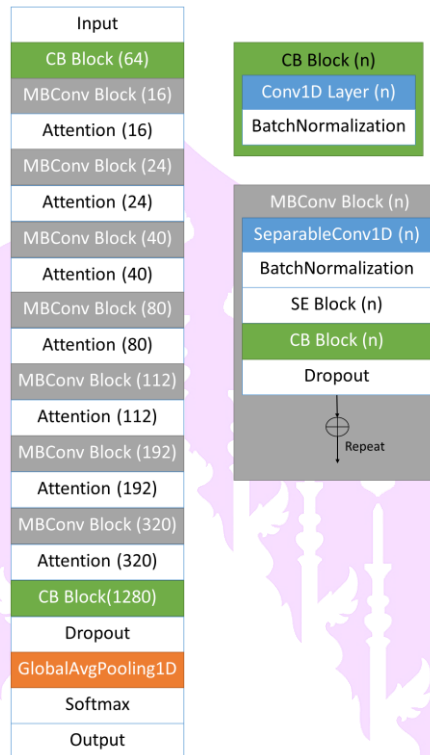
ภาพ 45 โครงสร้างแบบจำลอง Inception-V3

## 6) Xception



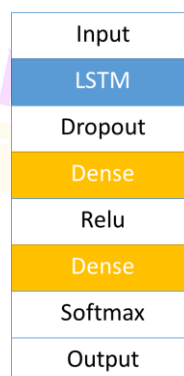
ภาพ 46 โครงสร้างแบบจำลอง Xception

## 7) EfficientNet-B0



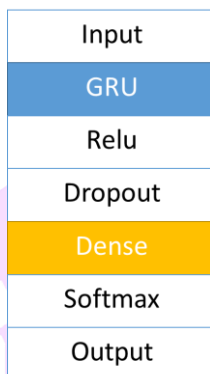
ภาพ 47 โครงสร้างแบบจำลอง EffcientNet-B0

## 8) LSTM model



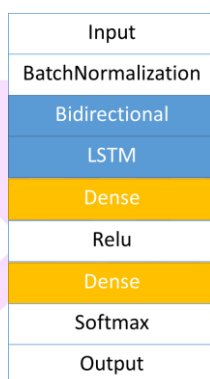
ภาพ 48 โครงสร้างแบบจำลอง LSTM

## 9) GRU model



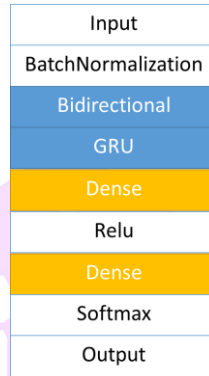
ภาพ 49 โครงสร้างแบบจำลอง GRU

## 10) Bi-LSTM model



ภาพ 50 โครงสร้างแบบจำลอง Bi-LSTM

## 11) Bi-GRU model



ภาพ 51 โครงสร้างแบบจำลอง Bi-GRU



## บทที่ 4

### ผลการวิจัย

การศึกษาค้นคว้าครั้งนี้ได้เปรียบเทียบความแม่นยำของอัลกอริทึมที่ใช้แบบจำลองหลายแบบกับชุดข้อมูลสาธารณะการรู้จำกิจกรรมของมนุษย์หลายชุดข้อมูล ในการทดลองครั้งนี้ได้ดำเนินการบน Google Colab Pro Tesla V100 พัฒนาด้วยภาษาไพทอน โดยใช้ไลบรารี TensorFlow, Keras, Scikit-Learn และ Numpy การทดสอบมีขั้นตอน ดังภาพ 52



ภาพ 52 แสดงภาพรวมการทำงานสำหรับเรียนรู้จำกิจกรรมในชีวิตประจำวันของมนุษย์

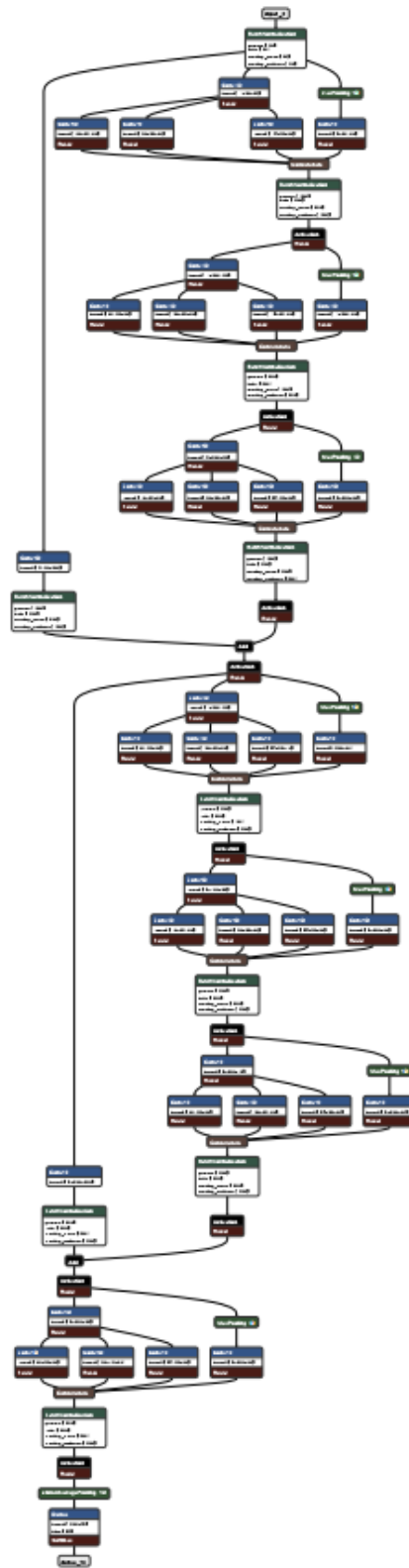
การทดสอบแรกเป็นการทดสอบกับชุดข้อมูลสาธารณะ PAMAP2 และเปรียบเทียบกับแบบจำลอง 7 แบบ ดังนี้ CNN LSTM BiLSTM GRU BiGRU CNN-LSTM และ InceptionTime ชุดข้อมูล PAMAP2 การประมวลผลเบื้องต้นใช้ median filter และ Butterworth แบบ Low-pass filter ลำดับที่ 3 หลังจากนั้นใช้วิธีการ Min-Max และ segmentation ที่ 2.56 วินาที 50% overlap โครงสร้างของแต่ละแบบจำลองแสดงดังตาราง 9 และ ภาพ 53

ตาราง 9 โครงสร้างแบบจำลอง 6 แบบกับชุดข้อมูล PAMAP2

Model	Layers
CNN	Conv1D — Dropout — Maxpooling — Flatten — Dense — Softmax
LSTM	LSTM — Dropout — Dense — Softmax
BiLSTM	BiLSTM — Dropout — Dense — Softmax
GRU	GRU — Dropout — Dense — Softmax
BiGRU	BiGRU — Dropout — Dense — Softmax
CNN-LSTM	Conv1D — Dropout — Maxpooling — Flatten — LSTM — Dropout — Dense — Softmax

ตาราง 10 ผลการเปรียบเทียบกับแบบจำลองทั้ง 7 แบบกับชุดข้อมูล PAMAP2

Model	Accuracy	Precision	Recall	F1-score	Loss	Parameters
CNN	86	87	86	86	0.622	836,983
LSTM	85	87	85	84	0.678	98,635
BiLSTM	81	70	97	81	0.688	195,915
GRU	86	88	86	85	0.578	77,899
BiGRU	84	86	84	83	0.493	154,443
CNN-LSTM	84	87	84	83	0.815	4,209,323
InceptionTime	88	74	99	85	0.584	1,943,835



ภาพ 53 โครงสร้างแบบจำลอง InceptionTime

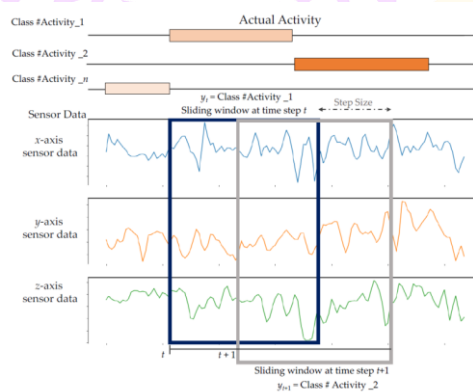
จากตาราง 10 แบบจำลอง InceptionTime ที่มีโครงสร้างเป็นแบบจำลองเป็นโครงข่ายประสาทเทียมให้ค่าความแม่นยำที่สูง แต่แบบจำลองมีความซับซ้อน พารามิเตอร์มีจำนวนมาก และใช้เวลาในการฝึกฝนนาน แต่แบบจำลอง LSTM Based ค่าความแม่นยำอยู่ในเกณฑ์ที่ยอมรับได้ และมีความซับซ้อนน้อยกว่าถึง 10 เท่าเมื่อเทียบกับจำนวนพารามิเตอร์ของ GRU ซึ่งเป็น LSTM Based ที่น้อยที่สุด แบบจำลอง LSTM Based ค่าความแม่นยำไม่ได้สูงเท่า CNN แต่ความแม่นยำอยู่ในเกณฑ์ที่ยอมรับได้ ใช้พารามิเตอร์น้อย และโครงสร้างของแบบจำลองซับซ้อนน้อยกว่า

ตาราง 11 ผลการเปรียบเทียบแต่ละกิจกรรมกับแบบจำลองทั้ง 6 แบบกับชุดข้อมูล PAMAP2

Activity	CNN	LSTM	BiLSTM	GRU	BiGRU	CNN-LSTM	Inception Time
Lying	0.95	0.92	0.95	0.96	0.94	0.96	0.97
Sitting	0.91	0.94	0.80	0.91	0.91	0.91	0.94
Standing	0.50	0.41	0.40	0.42	0.41	0.45	0.45
Walking	0.97	0.97	0.92	0.95	0.97	0.95	0.95
Running	0.96	0.96	0.96	0.97	0.95	0.94	0.96
Cycling	0.97	0.96	0.98	0.97	0.98	0.93	0.95
Nordic walking	0.98	0.96	0.91	0.97	0.98	0.95	1.00
Ascending stairs	0.81	0.76	0.63	0.73	0.67	0.82	0.81
Descending stairs	0.66	0.73	0.72	0.84	0.73	0.61	0.84
Vacuum cleaning	0.75	0.76	0.61	0.79	0.67	0.60	0.91
Ironing	0.86	0.80	0.81	0.79	0.81	0.84	0.81
Average	0.85	0.83	0.79	0.84	0.82	0.81	0.88

จากตาราง 11 แสดงผลการทดสอบแต่ละคลาสของกิจกรรมในชุดข้อมูล PAMAP2 กับแบบจำลองทั้ง 6 แบบที่กล่าวไปข้างต้น CNN กิจกรรมที่มีค่าความแม่นยำสูงได้แก่ การเดิน แบบนอร์ดิก กิจกรรมที่มีค่าความแม่นยำต่ำได้แก่ การยืน LSTM กิจกรรมที่มีค่าความแม่นยำสูงได้แก่ การเดิน กิจกรรมที่มีค่าความแม่นยำต่ำได้แก่ การเดินลงบันได BiLSTM กิจกรรมที่มีค่าความแม่นยำสูงได้แก่ การปั่นจักรยาน กิจกรรมที่มีค่าความแม่นยำต่ำได้แก่ การยืน GRU กิจกรรมที่มีค่าความแม่นยำสูงได้แก่ การวิ่ง การปั่นจักรยาน และการเดินแบบนอร์ดิก กิจกรรมที่มีค่าความแม่นยำต่ำได้แก่ การยืน BiGRU กิจกรรมที่มีค่าความแม่นยำสูงได้แก่ การปั่นจักรยาน และการเดินแบบนอร์ดิก กิจกรรมที่มีค่าความแม่นยำต่ำได้แก่ การยืน CNN-LSTM กิจกรรมที่มีค่าความแม่นยำสูงได้แก่ การอยู่นิ่ง กิจกรรมที่มีค่าความแม่นยำต่ำได้แก่ การยืน InceptionTime กิจกรรมที่มีค่าความแม่นยำสูงได้แก่ การเดินแบบนอร์ดิก กิจกรรมที่มีค่าความแม่นยำต่ำได้แก่ การยืน ผลลัพธ์แสดงให้เห็นว่าค่าความแม่นยำของแต่ละกิจกรรมไม่ได้ห่างกันจนเกินไป ค่าความแม่นยำโดยเฉลี่ยใกล้เคียงกัน จึงสรุปได้ว่าแบบจำลอง CNN Based มีความแม่นยำสูง แต่ใช้เวลาในการเทรนนาน แบบจำลองมีความซับซ้อนมาก และกินพลังงาน แต่ LSTM Based low power และมีความแม่นยำที่ยอมรับได้ แบบจำลองซับซ้อนน้อยกว่า CNN

การทดสอบที่สอผู้วิจัยได้ทดสอบกับ CNN Base ตัวอื่น ๆ ที่เป็นแบบจำลองน้ำหนักเบา (Lightweight) ดังนี้ VGG16, ResNet18, PyramidNet18, Inception-V3, Xception และ EfficientNet-B0 ในครั้งนี้ผู้วิจัยได้ทดสอบกับชุดข้อมูลสาธารณะ SPARS9x ขั้นตอนการประมวลผลเบื้องต้นใช้ median filter และ Butterworth แบบ Low-pass filter ลำดับที่ 3 การแบ่งข้อมูลใช้วิธีการ sliding window ที่ 4 วินาที 50% overlap แสดงดังภาพ 54

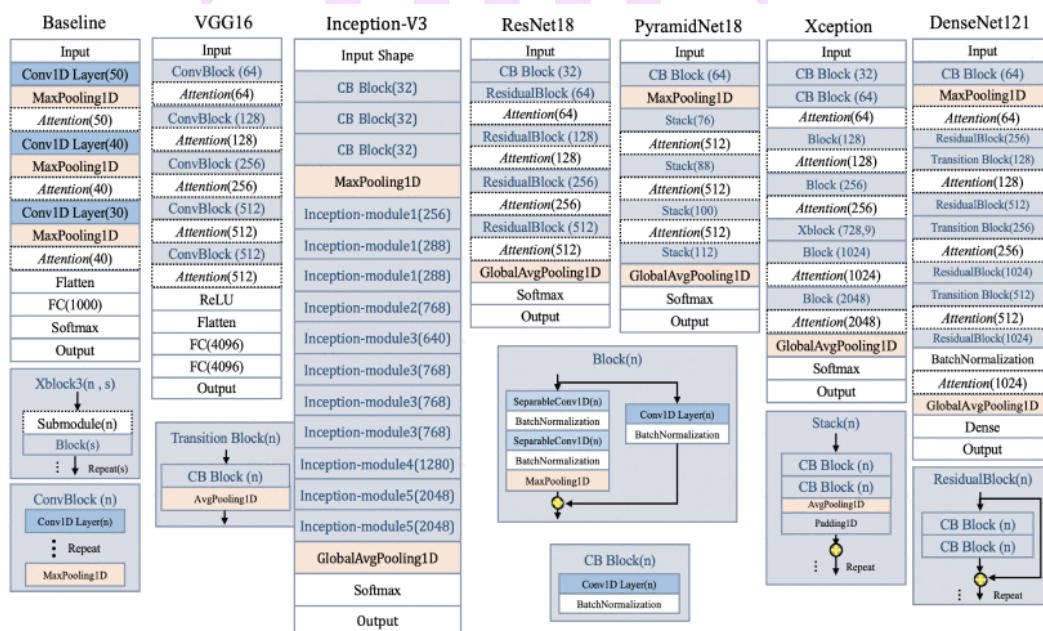


ภาพ 54 fix-width sliding window

ตาราง 12 ผลการเปรียบเทียบแบบจำลองโครงข่ายประสาทเทียม 6 แบบ กับชุดข้อมูล SPARS9x

Model	Accuracy	Loss	F1-score
VGG16	98.45% ( $\pm 0.310\%$ )	0.22 ( $\pm 0.082$ )	98.49% ( $\pm 0.310\%$ )
ResNet18	99.05% ( $\pm 0.317\%$ )	0.07 ( $\pm 0.010$ )	99.07% ( $\pm 0.316\%$ )
PyramidNet18	99.15% ( $\pm 0.379\%$ )	0.04 ( $\pm 0.021$ )	99.15% ( $\pm 0.380\%$ )
Inception-V3	98.84% ( $\pm 0.395\%$ )	0.05 ( $\pm 0.023$ )	98.86% ( $\pm 0.383\%$ )
Xception	99.07% ( $\pm 0.336\%$ )	0.04 ( $\pm 0.016$ )	99.09% ( $\pm 0.333\%$ )
EfficientNet-B0	98.43% ( $\pm 0.451\%$ )	0.09 ( $\pm 0.033$ )	98.45% ( $\pm 0.470\%$ )

จากตาราง 12 PyramidNet18 ให้ความแม่นยำ 99.15 ซึ่งมีความแม่นยำที่สูงเมื่อเทียบกับแบบจำลองที่นำมาเปรียบเทียบ จากผลการทดสอบก่อนหน้านี้ของ Inception มีความแม่นยำสูง ใช้พารามิเตอร์จำนวนมาก แต่ PyramidNet มีความซับซ้อนน้อยกว่า Inception อย่างเห็นได้ชัด แสดงดังภาพ 55



ภาพ 55 โครงสร้างแบบจำลอง Lightweight (Z. Zhongkai, 2022)

Backbone Model	Params	Layers	HASC	UCI	WISDM
Baseline	1.3M	10	87.71%	89.98%	89.55%
VGG16	<b>38.5M</b>	23	89.54%	90.40%	89.32%
Inception-V3	14.3M	313	91.85%	<b>94.23%</b>	<b>91.54%</b>
ResNet18	3.9M	67	90.53%	91.44%	87.53%
PyramidNet18	0.4M	74	91.48%	92.56%	85.23%
Xception	20.7M	126	92.31%	92.32%	90.17%
DenseNet121	5.6M	429	<b>92.55%</b>	92.52%	88.75%
MobileNet	6.0M	92	91.22%	80.32%	88.82%
MobileNet-V2	6.7M	156	90.62%	81.62%	74.71%
MobileNet-V3Small	2.9M	248	91.45%	91.25%	82.47%
NASNet Mobile	4.1M	<b>800</b>	86.46%	54.96%	84.98%
MnasNet	9.4M	147	89.75%	87.43%	84.57%
EfficientNet B0	11.4M	233	92.50%	93.53%	89.11%
EfficientNet lite0	10.8M	153	91.52%	91.20%	85.81%

ภาพ 56 จำนวนพารามิเตอร์ของแบบจำลอง Lightweight (Z. Zhongkai, 2022)

จากภาพ 55 และ 56 เป็นงานวิจัยของคุณ Zhao Zhongkai และคณะ แสดงให้เห็นว่าแบบจำลอง pyramidNet มีความซับซ้อนที่น้อยกว่า inception จากผลการทดสอบกับ CNN based สรุปได้ว่า CNN มีความแม่นยำในการรู้จำกิจกรรมมนุษย์สูง แต่แบบจำลองมีความซับซ้อนมาก และใช้จำนวนพารามิเตอร์ที่เยอะ

ตาราง 13 ผลการเปรียบเทียบกับแบบจำลองทั้ง 5 แบบกับชุดข้อมูล w-HAR

Model	Accuracy	Precision	Recall	F1-score
CNN	88.614%	88.538%	94.839%	88.319%
LSTM	95.125%	95.173%	95.125%	95.133%
BiLSTM	94.805%	94.839%	94.805%	94.803%
GRU	95.160%	95.167%	95.160%	95.158%
BiGRU	94.732%	94.755%	94.732%	94.734%

ตาราง 14 จำนวนพารามิเตอร์ของแบบจำลองทั้ง 5 แบบกับชุดข้อมูล w-HAR

Model	Parameters
CNN	2,046,215
LSTM	99,455
BiLSTM	197,119
GRU	69,639
BiGRU	138,247

ถัดมาผู้วิจัยได้นำ CNN มาทดสอบกับ LSTM Based และใช้ชุดข้อมูล w-HAR ซึ่งเป็นชุดข้อมูล low-power iot device โดยใช้แบบจำลอง CNN LSTM BiLSTM GRU และ BiGRU ทดสอบบนชุดข้อมูล w-HAR การประมวลผลเบื้องต้นใช้ median filter และ Butterworth แบบ Low-pass filter ลำดับที่ 3 การแบ่งข้อมูลใช้วิธีการ sliding window ที่ 3 วินาที 50% overlap

จากตาราง 13 แสดงให้เห็นว่าแบบจำลอง GRU ให้ค่าความแม่นยำอยู่ที่ 95.160% ซึ่งสูงกว่าแบบจำลองอื่น ๆ ที่นำมาเปรียบเทียบกับชุดข้อมูล w-HAR นอกจากนี้ค่าความแม่นยำที่สูงแล้วยังใช้จำนวนพารามิเตอร์ที่น้อยกว่าแบบจำลองอื่น ๆ แสดงดังตาราง 14 จะเห็นได้ว่า LSTM based มีความซับซ้อนน้อยกว่า CNN โดย GRU มีความซับซ้อนน้อยที่สุด ใช้พารามิเตอร์เพียง 69,639

การทดสอบสุดท้ายเป็นการทดสอบกับชุดข้อมูลสาธารณะ MEx กับแบบจำลอง CNN LSTM และ CNN-LSTM การประมวลผลเบื้องต้นใช้ averaged smoothing filters เพื่อลบสัญญาณรบกวนออก การแบ่งข้อมูลใช้วิธีการ sliding window ที่ 5 วินาที 50% overlap

ตาราง 15 ผลการเปรียบเทียบกับแบบจำลอง CNN LSTM CNN-LSTM กับชุดข้อมูล MEx ใช้เซ็นเซอร์วัดความเร่งบนตำแหน่งต้นขา

Model	Accuracy	Loss	F1-score
CNN	94.79% (+/- 1.205%)	0.25% (+/- 0.039%)	94.72% (+/- 1.262%)
LSTM	60.92% (+/- 3.862%)	0.98% (+/- 0.161%)	59.75% (+/- 3.688%)
CNN-LSTM	97.23% (+/- 0.524%)	0.13% (+/- 0.025%)	97.20% (+/- 0.534%)

ตาราง 16 ผลการเปรียบเทียบกับแบบจำลอง CNN LSTM CNN-LSTM กับชุดข้อมูล MEx ใช้เซ็นเซอร์วัดความเร่งบนตำแหน่งข้อมือ

Model	Accuracy	Loss	F1-score
CNN	80.69% (+/- 3.099%)	1.000% (+/- 0.238%)	80.49% (+/- 3.134%)
LSTM	43.42% (+/- 1.230%)	1.230% (+/- 0.235%)	41.34% (+/- 11.640%)
CNN-LSTM	91.72% (+/- 1.515%)	0.324% (+/- 0.040%)	91.69% (+/- 1.524%)

จากตาราง 15 และ 16 เป็นผลการทดสอบของแบบจำลอง CNN LSTM CNN-LSTM กับชุดข้อมูล MEx ใช้เซ็นเซอร์วัดความเร่งบนตำแหน่งต้นขา และข้อมือตามลำดับ ผลการทดสอบแบบจำลอง LSTM มีความแม่นยำน้อย แต่เมื่อนำ CNN มาพัฒนาร่วมด้วยส่งผลให้แบบจำลองมีความแม่นยำเพิ่มขึ้นมาจากเดิม



## บทที่ 5

### สรุปผลการวิจัย

จากการศึกษาและพัฒนาแบบจำลองการเรียนรู้เชิงลึกน้ำหนักเบาเพื่อการรู้จำกิจกรรมของมนุษย์โดยอาศัยข้อมูลเซนเซอร์จากสมาร์ตโฟน ซึ่งการทดลองนี้ได้มีวัตถุประสงค์เพื่อทำการพัฒนาอัลกอริทึมที่น้ำหนักเบาในการเรียนรู้เชิงลึกสำหรับการรู้จำกิจกรรมในชีวิตประจำวันของมนุษย์ที่มีประสิทธิภาพ และเพื่อนำแบบจำลองอัลกอริทึมที่ได้ ไปส่งเสริมพัฒนานวัตกรรมในการสนับสนุนดูแลมนุษย์ในด้านต่าง ๆ ต่อไปได้ โดยได้กำหนดขอบเขตของการวิจัยเป็นลำดับดังนี้ นำเสนอแบบจำลองสำหรับขั้นตอนการประมวลผลเบื้องต้น (data preprocessing) ขั้นตอนการเตรียมข้อมูล (data generation) และขั้นตอนการเรียนรู้จำกิจกรรมในชีวิตประจำวันของมนุษย์ (model training) ที่มีประสิทธิภาพ โดยได้ทำการทดลองทุกขั้นตอนทั้งในส่วนของการประมวลผลเบื้องต้น เป็นการนำชุดข้อมูลสาธารณะมาวิเคราะห์ และประมวลผลเพื่อจัดการกับข้อมูลให้พร้อมสำหรับขั้นตอนถัดไป ในขั้นตอนนี้ทางผู้วิจัยได้ทดลองเทคนิคต่าง ๆ ที่ได้ศึกษาค้นคว้ามา พบว่าเทคนิคต่าง ๆ ที่นำมาพัฒนาส่งผลต่อการฝึกฝนแบบจำลอง รวมถึงแต่ละเทคนิคแตกต่างกันจึงทำให้ขึ้นอยู่กับชุดข้อมูลนั้น ๆ โดยเทคนิคที่ผู้วิจัยนำมาพัฒนามีดังนี้ Min-Max Normalization, Butterworth แบบ Low-pass filter ลำดับที่ 3 และ segmentation ถัดมาเป็นขั้นตอนการเตรียมข้อมูลเพื่อนำเข้าแบบจำลอง ขั้นตอนนี้ผู้วิจัยได้ทำการแบ่งข้อมูลด้วยเทคนิค K-Fold Cross Validation และในขั้นตอนสุดท้ายการพัฒนาอัลกอริทึมเชิงลึกน้ำหนักเบา

จากการทดลองแรกทดสอบกับชุดข้อมูล PAMAP2 และแบบจำลอง CNN LSTM BiLSTM GRU BiGRU และ InceptionTime แบบจำลอง CNN Based มีความแม่นยำสูง แต่ใช้เวลาในการเทรนนาน แบบจำลองมีความซับซ้อนมาก และกินพลังงาน แต่ LSTM Based low power และมีความแม่นยำที่ยอมรับได้ แบบจำลองซับซ้อนน้อยกว่า CNN บนชุดข้อมูลสาธารณะ PAMAP2

การทดสอบต่อมาทดสอบกับแบบจำลองโครงข่ายประสาทเทียมน้ำหนักเบา ดังนี้ VGG16, ResNet18, PyramidNet18, Inception-V3, Xception และ EfficientNet-B0 บนชุดข้อมูลสาธารณะ SPARS9X ผลการทดสอบแบบจำลอง PyramidNet มีความซับซ้อนน้อยกว่า

Inception อย่างเห็นได้ชัดบนชุดข้อมูล SPARS9X โดยให้ค่าความถูกต้องอยู่ที่ 99.15% ค่าความสูญเสียโดยเฉลี่ยที่ 0.04% ค่าความสามารถของแบบจำลองโดยเฉลี่ยอยู่ที่ 99.15%

การทดสอบถัดมาเป็นการทดสอบบนชุดข้อมูล w-HAR ซึ่งเป็นชุดข้อมูล low-power iot device พบว่าแบบจำลอง LSTM based ให้ค่าความแม่นยำที่สูงกว่าแบบจำลอง CNN โดย GRU ให้ค่าความแม่นยำสูงบนชุดข้อมูล w-HAR

การทดสอบสุดท้ายได้ทดสอบกับชุดข้อมูล MEX กับแบบจำลอง CNN LSTM และ CNN-LSTM ผลการทดสอบแบบจำลอง LSTM มีความแม่นยำน้อย แต่เมื่อนำ CNN มาพัฒนา ร่วมกับส่งผลให้แบบจำลองมีความแม่นยำเพิ่มขึ้นมาจากเดิม

### การประยุกต์ผลการวิจัย

จากการศึกษาและวิจัยแสดงให้เห็นว่าอัลกอริทึมการเรียนรู้เชิงลึกนั้นสามารถขยายขอบเขตการทำงานได้อย่างกว้างขวางโดยเฉพาะทางด้านการเรียนรู้ของเครื่อง งานวิจัยนี้ได้นำมาประยุกต์ใช้ในการวิเคราะห์ข้อมูลที่อยู่ในรูปแบบของตัวเลข และข้อความจากค่าของเซนเซอร์ต่าง ๆ และข้อมูลของผู้เข้าร่วมในแต่ละชุดข้อมูล มาเพื่อใช้จำแนกกิจกรรมต่าง ๆ ในชีวิตประจำวันของมนุษย์ ซึ่งสามารถนำไปพัฒนาและประยุกต์ใช้ร่วมกับสมาร์ตโฟน หรืออุปกรณ์สวมใส่ที่ใช้เซนเซอร์ประเภทเดียวกันกับแบบจำลองได้

### ข้อเสนอแนะ

จากการทดลองพบว่า การพัฒนาแบบจำลองเชิงลึกนำหน้าหน้านั้น ไม่เพียงแต่การเลือกใช้อัลกอริทึมที่เหมาะสมเพื่อสร้างแบบจำลอง แต่ยังรวมไปถึงการวิเคราะห์ข้อมูลเบื้องต้นในขั้นตอนการประมวลผลข้อมูลเบื้องต้น ตลอดจนการเตรียมข้อมูล โดยในแต่ละขั้นตอนต้องใช้วิธีการในการวิเคราะห์และพัฒนาให้เหมาะสมกับประเภทและรูปแบบของข้อมูลที่จะวิเคราะห์

## บรรณานุกรม

- Agarwal Preeti และ Alam Mansaf. (2020). A lightweight deep learning model for human activity recognition on edge devices. 167, 2364-2373.
- Andrej Karpathy, Justin Johnson และ Li Fei-Fei. (2015). Visualizing and understanding recurrent networks. 1-11.
- Anguita Davide, Ghio Alessandro, Oneto Luca, Parra Xavier และ Reyes-Ortiz Jorge L. (2012). Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In **International workshop on ambient assisted living** (pp 216-223).
- Avilés-Cruz Carlos, Ferreyra-Ramírez Andrés, Zúñiga-López Arturo และ Villegas-Cortéz Juan. (2019). Coarse-fine convolutional deep-learning strategy for human activity recognition. 19(7), 1556.
- Bächlin Marc, Plotnik Meir, Roggen Daniel, Giladi Nir, Hausdorff Jeffrey M และ Tröster G. (2010). A wearable system to assist walking of Parkinson s disease patients. 49(01), 88-95.
- Baibich, M. N. Broto, J. M. Fert, A. Van Dau, F. N. Petroff, F. Etienne, P. และคณะ. (1988). Giant magnetoresistance of (001) Fe/(001) Cr magnetic superlattices. 61(21), 2472.
- Bock Marius, Hölzemann Alexander, Moeller Michael และ Van Laerhoven Kristof. (2021). Improving Deep Learning for HAR with Shallow LSTMs. In **2021 International Symposium on Wearable Computers** (pp 7-12).
- Dean Jeffrey Corrado Greg Monga Rajat Chen Kai Devin Matthieu Mao Mark และคณะ. (2012). Large scale distributed deep networks. 25.
- Eminoglu, B. (2017). **High performance FM gyroscopes** (พิมพ์ครั้งที่): University of California, Berkeley.
- Ganapati, B., Nicholas, T., Shill Holly และ Ogras, U. Y. (2020). w-HAR: An activity recognition dataset and framework using low-power wearable devices. 20(18), 5356.
- Garcia-Gonzalez, D., Rivero, D., Fernandez-Blanco, E. และ Luaces, M. R. (2020). A public domain dataset for real-life human activity recognition using smartphone sensors. 20(8), 2200.
- Grahn, E. (2017). Evaluation of MEMS accelerometer and gyroscope for orientation tracking nutrunner functionality.
- He Kaiming, Zhang Xiangyu, Ren Shaoqing และ Sun Jian. (2016). Deep residual learning for image

- recognition. In **Proceedings of the IEEE conference on computer vision and pattern recognition** (pp 770-778).
- Hecht-Nielsen Robert. (1992). **Theory of the backpropagation neural network**: Elsevier.
- Hinton Geoffrey E, Osindero Simon and Teh Yee-Whye. (2006). A fast learning algorithm for deep belief nets. 18(7), 1527-1554.
- Ignatov Andrey. (2018). Real-time human activity recognition from accelerometer data using Convolutional Neural Networks. 62, 915-922.
- Micucci Daniela , Mobilio Marco and Napolitano Paolo. (2017). Unimib shar: A dataset for human activity recognition using acceleration data from smartphones. 7(10), 1101.
- Mohamadabadi, K. (2013). **Anisotropic Magneto-resistance Magnetometer for inertial navigation systems**. Ecole Polytechnique X
- Ordóñez Francisco Javier and Roggen Daniel. (2016). Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. 16(1), 115.
- Philip Boyer, Burns., D. and Cari Whyne. (2020). Shoulder Physiotherapy Activity Recognition 9-Axis Dataset. IEEE Dataport. doi:10.21227/cx5v-vw46.
- Prandi, L. Caminada, C. Coronato, L. Cazzaniga, G. Biganzoli, F. Antonello, R. and Kijum. (2011). A low-power 3-axis digital-output MEMS gyroscope with single drive and multiplexed angular rate readout. In **2011 IEEE International Solid-State Circuits Conference** (pp 104-106).
- Reiss Attila and Stricker Didier. (2012). Introducing a new benchmarked dataset for activity monitoring. In **2012 16th international symposium on wearable computers** (pp 108-109).
- Roggen Daniel Calatroni Alberto Rossi Mirco Holleczeck Thomas Förster Kilian Tröster Gerhard and Kijum. (2010). Collecting complex activity datasets in highly rich networked sensor environments. In **2010 Seventh international conference on networked sensing systems (INSS)** (pp 233-240).
- Ronald Mutegeki, Poulouse Alwin and Han Dong Seog. (2021). iSPLInception: an inception-ResNet deep learning architecture for human activity recognition. 9, 68985-69001.
- Scholl Philipp M, Wille Matthias and Van Laerhoven Kristof. (2015). Wearables in the wet lab: a laboratory system for capturing and guiding experiments. In **Proceedings of the 2015**

- ACM International Joint Conference on Pervasive and Ubiquitous Computing** (pp 589-599).
- Selany, S., Nasution, S. M. และ Purboyo, T. W. (2015). Sosemanuk Algorithm for Encryption and Decryption Video on Demand (VoD). In **2015 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)** (pp 125-129).
- Shryock, H. S. (2004). **The Methods and Materials of Demography** (พิมพ์ครั้งที่ 1). New York: Academic Press.
- Stisen Allan Blunck Henrik Bhattacharya Sourav Prentow Thor Siiger Kjærgaard Mikkel Baun Dey Anind และคณะ. (2015). Smart devices are different: Assessing and mitigating mobile sensing heterogeneities for activity recognition. In **Proceedings of the 13th ACM conference on embedded networked sensor systems** (pp 127-140).
- Sugara Nashar Luthfi, Purboyo Tito Waluyo และ Prasasti Anggunmeka Luhur. (2018). Comparative analysis of image compression using huffman and DCT Method on JPG Image. 13, 4447-4452.
- Szegedy Christian, Ioffe Sergey, Vanhoucke Vincent และ Alemi Alexander A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In **Thirty-first AAAI conference on artificial intelligence** (pp. 1-9).
- Sztyley Timo และ Stuckenschmidt Heiner. (2016). On-body localization of wearable devices: An investigation of position-aware activity recognition. In **2016 IEEE International Conference on Pervasive Computing and Communications (PerCom)** (pp 1-9).
- Wang, J., Chen, Y., Hao, S., Peng, X. และ Hu, L. J. P. r. l. (2019). Deep learning for sensor-based activity recognition: A survey. 119, 3-11.
- Wang LuKun และ Liu RuYue. (2020). Human activity recognition based on wearable sensor using hierarchical deep LSTM networks. 39(2), 837-856.
- Weiss Gary M, Yoneda Kenichi และ Hayajneh Thaier. (2019). Smartphone and smartwatch-based biometrics using activities of daily living. 7, 133190-133202.
- Wijekoon, A., Wiratunga, N. และ Cooper, K. J. a. p. a. (2019). Mex: Multi-modal exercises dataset for human activity recognition.
- Z. Zhongkai, S. K., K. Kondo, T. Hasegawa and M. Koshino. (2022). A Comparative Study: Toward an Effective Convolutional Neural Network Architecture for Sensor-Based Human

Activity Recognition. 10, 20547-20558. doi:10.1109/ACCESS.2022.3152530.

เฉลิมเกียรติ สอนเสียม (2560). In การประชุมวิชาการระดับชาติ ครั้งที่ 17และการประชุมวิชาการระดับนานาชาติ ครั้งที่ 7 มหาวิทยาลัยศรีปทุม (pp 2389). การประชุมออนไลน์. Retrieved 27 ตุลาคม 2565, from <https://spucon.spu.ac.th/>

ทีฆะ, ช. (2560). การประยุกต์ใช้เซนเซอร์ บนสมาร์ตโฟนในการจัดการเรียนการสอนฟิสิกส์ กลศาสตร์. In การประชุมวิชาการระดับชาติและนานาชาติมหาวิทยาลัยศรีปทุม ครั้งที่12 ประจำปี 2560 (หน้า 2388-2389). มหาวิทยาลัยศรีปทุม.

เอกภพ สกุกกิจกาญจน์. (2552). ระบบเซ็นเซอร์และออกงานด้วยลายนิ้วมือผ่านทางเว็บไซต์. มหาวิทยาลัยเทคโนโลยีมหานคร, กรุงเทพฯ.





## ประวัติผู้วิจัย

ชื่อ-สกุล	นางสาวพรนิภา จันทวงศ์
วัน เดือน ปี เกิด	6 มีนาคม 2540
สถานที่เกิด	โรงพยาบาลแพร์
วุฒิการศึกษา	วท.บ. เทคโนโลยีสารสนเทศ มหาวิทยาลัยพะเยา
ที่อยู่ปัจจุบัน	136 ม.3 ต.น้ำเลา อ.ร้องกวาง จ.แพร่ 54140
ผลงานตีพิมพ์	<p>P. Jantawong, S. Mekruksavanich and A. Jitpattanakul, "Monitoring System of Wearable Sensor Signal in Rehabilitation Using Efficient Deep Learning Approaches," 2022 26th International Computer Science and Engineering Conference (ICSEC), Sakon Nakhon, Thailand, 2022, pp. 361–365, doi: 10.1109/ICSEC56337.2022.10049326.</p> <p>P. Jantawong, N. Hnoohom, A. Jitpattanakul and S. Mekruksavanich, "Time Series Classification Using Deep Learning for HAR Based on Smart Wearable Sensors," 2022 26th International Computer Science and Engineering Conference (ICSEC), Sakon Nakhon, Thailand, 2022, pp. 357–360, doi: 10.1109/ICSEC56337.2022.10049357.</p> <p>Mekruksavanich, S., Jantawong, P., Hnoohom, N., &amp; Jitpattanakul, A. (2022). Recognizing Driver Activities Using Deep Learning Approaches Based on Smartphone Sensors. In International Conference on Multi-disciplinary Trends in Artificial Intelligence (pp. 146–155). Springer, Cham.</p> <p>Mekruksavanich, S., Jantawong, P., Hnoohom, N., &amp; Jitpattanakul, A. (2022). Wearable Fall Detection Based on Motion Signals Using Hybrid Deep Residual Neural Network. In International Conference on Multi-disciplinary Trends in Artificial Intelligence (pp. 216–224). Springer, Cham.</p> <p>Mekruksavanich, S., Jantawong, P., Hnoohom, N., &amp; Jitpattanakul,</p>

A. (2022, June). Deep Learning Models for Daily Living Activity Recognition based on Wearable Inertial Sensors. In 2022 19th International Joint Conference on Computer Science and Software Engineering (JCSSE) (pp. 1–5). IEEE.

Mekruksavanich, S., Jantawong, P., Hnoohom, N., & Jitpattanakul, A. (2022, October). Badminton Activity Recognition and Player Assessment based on Motion Signals using Deep Residual Network. In 2022 IEEE 13th International Conference on Software Engineering and Service Science (ICSESS) (pp. 80–83). IEEE.

Mekruksavanich, S., Jantawong, P., Hnoohom, N., & Jitpattanakul, A. (2022, October). Refined LSTM Network for Sensor-based Human Activity Recognition in Real World Scenario. In 2022 IEEE 13th International Conference on Software Engineering and Service Science (ICSESS) (pp. 256–259). IEEE.

Mekruksavanich, S., Jantawong, P., Hnoohom, N., & Jitpattanakul, A. (2022, September). Heterogeneous Recognition of Human Activity with CNN and RNN-based Networks using Smartphone and Smartwatch Sensors. In 2022 3rd International Conference on Big Data Analytics and Practices (IBDAP) (pp. 21–26). IEEE.

Mekruksavanich, S., Jantawong, P., Hnoohom, N., & Jitpattanakul, A. (2022, September). ResNet-based Network for Recognizing Daily and Transitional Activities based on Smartphone Sensors. In 2022 3rd International Conference on Big Data Analytics and Practices (IBDAP) (pp. 27–30). IEEE.

Mekruksavanich, S., Jantawong, P., Hnoohom, N., & Jitpattanakul, A. (2022, August). The Effect of Sensor Placement for Accurate Fall Detection based on Deep Learning Model. In 2022 Research, Invention, and Innovation Congress: Innovative Electricals and Electronics (RI2C) (pp. 124–129). IEEE.

Mekruksavanich, S., Jantawong, P., Hnoohom, N., & Jitpattanakul,

A. (2022, August). Hyperparameter Tuning in Convolutional Neural Network for Face Touching Activity Recognition using Accelerometer Data. In 2022 Research, Invention, and Innovation Congress: Innovative Electricals and Electronics (RI2C) (pp. 101–105). IEEE.

Mekruksavanich, S., Jantawong, P., Hnoohom, N., & Jitpattanakul, A. (2022, August). Deep Learning Networks for Eating and Drinking Recognition based on Smartwatch Sensors. In 2022 Research, Invention, and Innovation Congress: Innovative Electricals and Electronics (RI2C) (pp. 106–111). IEEE.

Mekruksavanich, S., Jantawong, P., Hnoohom, N., & Jitpattanakul, A. (2022, August). Classification of Physical Exercise Activity from ECG, PPG and IMU Sensors using Deep Residual Network. In 2022 Research, Invention, and Innovation Congress: Innovative Electricals and Electronics (RI2C) (pp. 130–134). IEEE.

Mekruksavanich, S., Jantawong, P., & Jitpattanakul, A. (2022, July). Deep Residual Networks for Human Activity Recognition based on Biosignals from Wearable Devices. In 2022 45th International Conference on Telecommunications and Signal Processing (TSP) (pp. 310–313). IEEE.

Mekruksavanich, S., Jantawong, P., & Jitpattanakul, A. (2022, January). Deep Learning-based Action Recognition for Pedestrian Indoor Localization using Smartphone Inertial Sensors. In 2022 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON) (pp. 346–349). IEEE.

Mekruksavanich, Sakorn & Jantawong, Ponnipa & Charoenphol, Adisaya & Jitpattanakul, Anuchit. (2021). Fall Detection from Smart Wearable Sensors using Deep Convolutional Neural Network with Squeeze-and-Excitation Module. 448–453.

10.1109/ICSEC53205.2021.9684626.

Jantawong, Ponnipa & Hnoohom, Narit & Jitpattanakul, Anuchit & Mekruksavanich, Sakorn. (2021). A Lightweight Deep Learning Network for Sensor-based Human Activity Recognition using IMU sensors of a Low-Power Wearable Device. 459–463.

10.1109/ICSEC53205.2021.9684631.

Mekruksavanich, Sakorn & Jantawong, Ponnipa & Hnoohom, Narit & Jitpattanakul, Anuchit. (2021). Bidirectional Gate Recurrent Unit Neural Network for Recognizing Face Touching Activities using Smartwatch Sensors. 454–458.

10.1109/ICSEC53205.2021.9684642.

Jantawong, Ponnipa & Jitpattanakul, Anuchit & Mekruksavanich, Sakorn. (2021). Enhancement of Human Complex Activity Recognition using Wearable Sensors Data with InceptionTime Network. 12–16. 10.1109/IBDAP52511.2021.9552133.

Mekruksavanich, Sakorn & Jantawong, Ponnipa & Jitpattanakul, Anuchit. (2021). A Lightweight Deep Convolutional Neural Network with Squeeze-and-Excitation Modules for Efficient Human Activity Recognition using Smartphone Sensors. 23–27.

10.1109/IBDAP52511.2021.9552111.